Evaluating the Safety and Effectiveness of Continuous Glucose Monitors for Glycaemic Control in Intensive Care

Tony Zhou

A thesis presented for the degree of

Doctor of Philosophy

in

Mechanical Engineering

at the

University of Canterbury,

Christchurch, New Zealand

September 2019

Acknowledgements

To my supervisor, Distinguished Professor Geoff Chase, thank you for providing me with the opportunity to undertake this study and pursue a PhD. It had been a goal of mine for a while, and I thank you for all the opportunities that you've given me, as well the guidance and support.

To my associate supervisor, Dr Jennifer Knopp, thank you also for the support and wisdom you've provided over the years. I will always be thankful for you telling me to stick it out for another month when I was about 8 months in and ready to throw in the towel because I didn't think I was good enough!

To Chris, and everyone in the bio-engineering office that has come through while I've been here. You all make the office a great place to be every day, and I am grateful for your banter and our shared adventures, both in New Zealand and beyond.

To all of the University of Canterbury Hockey Club members that I've had the pleasure of getting to know over the years, thank you for providing such a fun and competitive outlet outside of university. Particularly to those that I have played with for many years, including Nick and Ben, I hope I have the pleasure of sharing a turf with you again, and if not, then a beer!

To my sister Lisa, thank you for all your advice over the years. And my brother Will, thank you for your support through everything, I know you are always there for me. I look up to you both for your academic and sporting achievements, and try to emulate them as best I can.

And last but not least, thank you Mum and Dad, two of the hardest working people I know. Thank you for working in a Fish and Chips/Chinese Takeaways shop 12 hours a day, 6 days a week for the last 20 years to make sure your 3 kids from Titahi Bay went to university.

Contents

Abstract

Glycaemic control (GC) in the intensive care unit is clinically contentious. Hyperglycaemia, hypoglycaemia, and glycaemic variability are increased with many GC protocols, and, critically, all associated with increased morbidity and mortality. While some studies and physiological evidence suggests GC should benefit hyperglycaemic patients, others show no or negative effects and increased incidence of hypoglycaemia. Interpretation of results is made more difficult by differences in the measurement and reporting of glycaemic control, blood glucose (BG) levels and variability in patients. In addition, target ranges for glycaemic control are not universally accepted, and higher targets are often used out of fear of hypoglycaemia, rather than their relationship to a clinical outcome.

The issues of hyperglycaemia, hypoglycaemia, and glycaemic variability may be solved with the help of continuous glucose monitoring (CGM) devices. CGM sensors have a higher rate of BG measurement, and have been effective in managing diabetes, while offering potential benefits for use in the intensive care unit (ICU). Use of CGM devices in the ICU has been limited, primarily due to the higher point accuracy errors over currently used traditional intermittent blood glucose measures. General models of CGM errors, including drift and random errors, are lacking, but would enable better design of protocols to safely and effectively utilise these devices as integrated parts of GC protocols.

This research presents an auto-regressive (AR) based modelling method that separately characterises the drift and random noise of the GlySure CGM sensor (GlySure LLC, UK). Clinical sensor data (n=33) and reference measurements were used to generate 2 AR models to describe sensor drift and noise. These models were used to generate 100 Monte-Carlo simulations based on reference blood glucose measurements. These simulated CGM traces were then compared to the original CGM clinical data using mean absolute relative difference (MARD) and a Trend Compass. The point accuracy MARD was very similar between simulated and clinical data (9.6% vs 9.9%). A Trend Compass was used to

assess trend accuracy, and found simulated and clinical sensor profiles were similar (simulated trend index 11.4° vs. clinical trend index 10.9°).

The model and method accurately represents cohort sensor behaviour over patients, providing a general modelling approach to any such sensor by separately characterising each type of error that can arise in the data. Overall, it enables better protocol design using validated virtual patients based on accurate expected CGM sensor behaviour, as well as enabling the analysis of what level of each type of sensor error would be necessary to obtain desired glycaemic control safety and performance with a given protocol.

The modelling of CGM sensors may then be used for in-silico virtual trials to replace intermittent measurements in GC. This research aims to delineate the trade-offs of performance, safety and workload that CGM sensors provide in GC protocols. Clinical data from 236 patients were used for clinically validated virtual trials. A CGM-enabled version of the STAR GC protocol was used to evaluate the use of guard rails and rolling windows. Safety was assessed through percentage of patients who had a severe hypoglycaemic episode (BG < 2.22 mmol/L) as well as percentage of resampled BG < 4.0 mmol/L. Performance was assessed as percentage of resampled measurements in the 4.4-7.0 mmol/L and the 4.4-8.0 mmol/L target bands. Workload was measured by number of manual BG measures per day.

CGM-enabled versions of STAR decreased the number of required blood draws by up to 74%, while maintaining performance (76.6% BG measurements in the 4.4-7.0 mmol/L range vs. 62.8% clinically, 87.9% in the 4.4-8.0 mmol/L range vs. 83.7% clinically) and maintaining patient safety (1.13% of patients experienced a severe hypoglycaemic event vs. 0.85% clinically, 1.37% of BG measurements were less than 4.0 mmol/L vs. 0.51% clinically). CGM sensor traces were simulated in virtual trials and used in place of intermittent BG measurements to guide GC interventions and decisions. Existing GC protocols, such as STAR, may only need to be adjusted slightly to gain the benefits of the increased

temporal measurements of CGM sensors, through which workload may be significantly decreased while maintaining GC performance and safety.

This research also reviews differences in the reporting of BG level and its variability in literature, which is a growing issue with the emergence of CGM sensors and other high rate sensors. There are already a multitude of differences in the reporting of BG level and variability where only intermittent measurements are concerned. The rise of new high rate sensing technology can add more temporal factors that also need to be considered, but also adds to the differences of reporting of BG and variability. The research then proposes a vision for improved description of glycaemia and how it changes and evolves over time. This work then presents a continuous glucose monitoring sensor-based method to better quantify glycaemic level and variability, based on clinically defined metrics. A case study of this new method is presented using CGM sensor data from a study of 614 infants at risk of neonatal hypoglycaemia. The CGM sensor data is used to understand glycaemia and how it evolves over time in an infant cohort. Results show the new clinically defined method is able to describe changes in glycaemic level and variability in these patients and presents a flexible way forward for accurately describing state and variability from a clinically defined perspective. This method may provide better insight to patient glycaemia over time, and thus provide scope for improved control of glycaemia. The metric is then used to assess glycaemic State Changes in the infant cohort, and attempts to relate the number of State Changes per Day to neurodevelopmental impairment data from the original study from which this infant cohort data is provided (the CHYLD Study). State Changes per Day as a metric for variability alone was found to be a weak indicator of neurodevelopmental impairment, mostly due to the complex behaviour experienced by the infants recovering from hypoglycaemia identified as having impairment.

Overall, this thesis has approached the problem of using CGM sensors to understand glycaemia and its evolution over time, and to inform model-based, personalised GC. It develops and validates a novel and

accurate CGM sensor modelling method. It uses this methodology in redesigning a successful model-based GC protocol for optimised use with CGM sensors. Finally, it presents and validates a method of evaluating trends and variability in high rate CGM sensor data that results from CGM enabled GC. Thus, the thesis has developed and validated a series of solutions to CGM enabled GC.

Deputy Vice-Chancellor's Office
Postgraduate Research Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 4, CGM Modelling, is extracted from "Zhou, T., Dickson, J.L., and Geoffrey Chase, J., Autoregressive Modeling of Drift and Random Error to Characterize a Continuous Intravascular Glucose Monitoring Sensor. J Diabetes Sci Technol, 2018. 12(1): p. 90-104."

Please detail the nature and extent (%) of contribution by the candidate:

The candidate developed the hypothesis, necessary algorithms for analysis, conducted the analysis and wrote the article, contributing a total of 70%.

Co-authors contributions varied from contributing toward formation of high level hypothesis, supporting the candidate in the form of proof reading of the article and including feedback on the clarity of the candidate's results and discussion.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the Doctoral candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Tony Zhou* Signature: *Tony Zhou* Date: *13/09/19*

Deputy Vice-Chancellor's Office
Postgraduate Research Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 5, CGM STAR, is extracted from "Zhou, T., Dickson, J.L., Shaw, G.M., and Chase, J.G., Continuous Glucose Monitoring Measures Can Be Used for Glycemic Control in the ICU: An In-Silico Study. J Diabetes Sci Technol, 2017: p. 1932296817738791."

Please detail the nature and extent (%) of contribution by the candidate:

The candidate developed the hypothesis, necessary algorithms for analysis, conducted the analysis and wrote the article, contributing a total of 70%.

Co-authors contributions varied from contributing toward formation of high level hypothesis, supporting the candidate in the form of proof reading of the article and including feedback on the clarity of the candidate's results and discussion.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the Doctoral candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Tony Zhou* Signature: *Tony Zhou* Date: *13/09/19*

Deputy Vice-Chancellor's Office
Postgraduate Research Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 6, Variability and Level Analysis of CGM Sensor Traces, is extracted from "Zhou, T., Knopp, J.L., and Chase, J.G., The state of variability: A vision for descriptors of glycaemia. Annual Reviews in Control, 2019."

Please detail the nature and extent (%) of contribution by the candidate:

The candidate developed the hypothesis, necessary algorithms for analysis, conducted the analysis and wrote the article, contributing a total of 70%.

Co-authors contributions varied from contributing toward formation of high level hypothesis, supporting the candidate in the form of proof reading of the article and including feedback on the clarity of the candidate's results and discussion.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the Doctoral candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Tony Zhou* Signature: *Tony Zhou* Date: *13/09/19*

List of Figures

List of Tables

Nomenclature

ADRR	Average Daily Risk Range
AR	Auto-regressive
AUC	Area under curve
BG	Blood glucose
CEG	Clarke error grid
CGM	Continuous glucose monitoring
CHYLD	Children with Hypoglycaemia and their Later Development
CoV	Coefficient of Variability
d	Drift
GC	Glycaemic control
HBGI	High Blood Glucose Index
ICING	Intensive Care Insulin-Nutrition-Glucose
ICU	Intensive Care Unit
IG	Interstitial glucose
IQR	Inter-quartile range
IV	intravascular
IM	Intermittent
LBGI	Low Blood Glucose Index
MAD	Mean absolute difference
MAGE	Mean amplitude of glucose excursion
MARD	Mean absolute relative difference
MODD	Mean of Daily Differences
NICU	Neonatal intensive care unit
POC	Point-of-care
ROC	Receiver operator characteristic
SD	Standard deviation
SF	Sensorflux or Sensor fluctuations
SG	Sensor glucose
S_I	Insulin sensitivity
SPRINT	Specialised Relative Insulin-Nutrition Tables
STAR	Stochastic Targeted
TiB	Time in band

Chapter 1 Clinical Introduction

1.1 Physiology and Impact of Hyperglycaemia

Critically ill patients commonly experience stress induced insulin resistance resulting in hyperglycaemia [1-7]. Events such as major trauma and surgery cause counter-regulatory hormones such as cortisol, glucagon, catecholamines and growth hormone to be significantly elevated [8-11]. The elevation of these hormones cause increased hepatic glucose production, inhibition of insulin secretion and peripheral insulin resistance [11], resulting in increased blood glucose (BG) concentrations, even for patients with no prior history of diabetes.

Studies have shown that there is a strong association between hyperglycaemia and intensive care unit (ICU) mortality and morbidity [1, 3, 6, 12]. Hyperglycaemia can also lead to further complications including myocardial infarction [1], sepsis and infection [13-17], polyneuropathy, and multiple organ failure [2]. It also has been observed to induce oxidative stress and endothelial and microcirculation dysfunction [18], both of which contribute to organ failure [19, 20]. Thus, controlling glycaemia to normal levels may be key to reducing these negative outcomes in this clinical cohort.

Glycaemic variability alone has also been shown to result in negative clinical outcomes [21-26]. This glycaemic variability is a result of various factors, including intra-patient variability due to unpredictable patient state or response to care, and also can be a result of poor control causing erratic changes in BG. In addition, hypoglycaemia has also been associated with poor clinical outcomes, independently of glycaemic variability [23, 26-29]. Hypoglycaemia appears more directly detrimental to patients than hyperglycaemia [26]. As such, any control of level and variability of glycaemia needs to be carefully designed as to not decrease BG too much while it is already low by normal physiological standards, where GC, patient variability, and hypoglycaemia are all linked [21, 30-34].

1.2 Clinical Care and Control

Clinical studies have shown insulin therapy can improve patient outcomes [2, 7, 35, 36]. However, several studies failed to replicate early beneficial results [15, 37-39], while others showed beneficial outcomes [40, 41] and several showed no benefit or harm [30, 42, 43]. Overall, one large randomised trial in 2009 reporting adverse outcomes has become a point of reference for many rejecting glycaemic control (GC) [39, 44].

In theory, insulin therapy should directly counteract each of the mechanisms of injury associated with hyperglycaemia, reducing inflammation and infection, protecting endothelial cells and reducing oxidative stress. The physiological evidence suggests safe, effective glycaemic control should benefit all patients in the ICU through reduced blood glucose (BG) and the insulin hormone itself [5, 45-47]. However, for many of the studies, factors such as BG variability, inter- and intra-patient variability, and hypoglycaemia have made good GC hard to achieve and very difficult to deliver safely, effectively and consistently. These factors are prevalent in the studies for which intensive GC has resulted in worse outcomes than more traditional GC. They confound study outcomes and comparisons across studies, making it hard to deduce the optimal level and method of control.

Only one study, of all reported, reduced hypoglycaemia and improved outcomes with insulin therapy for GC [36], indicating poor control, rather than GC itself may be the cause of the variability in study outcomes. A recent study of metabolic variability and mortality outcomes showed no difference in underlying metabolic variability between survivors and non-survivors [32], suggesting patient outcomes are a function of the quality of control delivered and not patient-specific or cohort-specific characteristics. Poorly delivered control might thus affect study outcomes when looking for benefit or harm, as studies show good GC may need to be achieved for essentially all patients to show benefit [48]. As such, there is a clear need for good control algorithms able to directly manage the significant intra- and inter- patient variability that makes GC difficult [49-53] to provide safe, effective control for

all patients.

Model-based methods have been developed to achieve this goal [34, 41, 54-56]. In particular, model-based methods monitor and respond to changes in patient-specific metabolic condition, typically using an insulin sensitivity parameter. Insulin sensitivity is a key determinant of the glucose uptake response to an insulin dose, and this sensitivity is most variable early in the ICU stay where most hypoglycaemia occurs [23], both in response to patient condition and clinical interventions [24, 57-59]. Hence, it is a key factor in managing glycaemia and patient variability [50-53].

1.3 Glycaemic Control in the Christchurch Hospital ICU

The Specialised Relative Insulin Nutrition Tables (SPRINT) protocol was a model-derived paper-based GC protocol developed jointly by the Christchurch hospital ICU and the University of Canterbury Bioengineering department. The GC protocol was able to achieve good control with significantly reduced incidence of hypoglycaemia in comparison to the previous standards of practice within the Christchurch hospital ICU [36, 48, 60]. Under the SPRINT protocol reduced mortality and organ failure was observed, suggesting glycaemic control provides outcome benefits to ICU patients [48].

More recently, STAR (Stochastic TARgeted), a computerised model-based protocol utilising stochastic models to forecast glycaemic outcomes for a given insulin and nutrition intervention, has replaced SPRINT as the standard of care within the Christchurch hospital ICU. STAR was shown to reduce the incidence of hypoglycaemia, and increase nutrition rates in comparison to SPRINT [34, 61, 62]. STAR also reduced clinical workload, with the average number of BG measurements taken by a nurse in STAR reduced to 13.6 per day in comparison to 15.8 per day under SPRINT [34].

The STAR protocol has been further developed to create the STOMP (STOchastic Model Predictive) controller, a protocol similar to STAR that formalises the control methodology using model predictive control theory. The protocol utilises cost functions to determine optimal nutrition-insulin interventions and was able to reduce required BG measurement frequency while also increasing ease of controller tuning. The controller has been developed and tested *in-silico* and is awaiting clinical trials for further validation [63].

1.4 Challenges to Glycaemic Control

However, while the future looks optimistic for model-based GC and patient outcomes [44], there is still much debate over what GC targets are appropriate and/or safe for an ICU context [39]. Many clinicians prefer higher targets out of fear of hypoglycaemia [26-29]. In contrast, most studies which have shown benefit have shown it from using lower targets [25, 40, 64, 65]. Further, additional evidence suggests these targets may be patient specific and/or differ between ICU cohorts [66]. Confounding this issue is a lack of consensus on how to measure and/or report GC outcomes and variability at a cohort and patient level [30, 67].

Workload in the ICU also has a significant impact on the clinical efficacy and compliance of glycaemic control. Many studies and pilot trials have reported the extra clinical effort required to implement intensive insulin therapy in ICUs [68-74]. This issue is especially valid in ICUs where a single nurse is responsible for multiple patients simultaneously [68]. Clinical incompliance to a particular GC protocol may result in a reduction in performance and safety, resulting in more adverse outcomes. Thus, nurse workload often has a direct effect on protocol compliance, and thus GC performance and clinical outcomes, where labour intensive but otherwise 'good' GC protocols are unlikely to be clinical practicable or beneficial.

Some researchers believe an upcoming technology, continuous glucose monitoring (CGM) sensors, would solve many of these issues of measuring and mitigating both variability and hypoglycaemia, and also decreasing nurse workload. CGM sensors have a far higher measurement frequency than the intermittent point-of-care (POC) BG measurements, which are the current standard for GC. This higher measurement frequency can allow better identification and quantification of variability in BG, as well as the ability to identify hypoglycaemia much faster than intermittent measurements. Both of the factors would, in theory, combine to make GC easier. However, difficulties in control still remain, and CGM sensor technology has not yet been shown to solve all of the issues of measurement and control that researchers would have liked.

1.5 Preface

This chapter provides the background for the clinical problem addressed here. Specifically, this thesis evaluates the advantages and disadvantages of using CGM sensors for GC in the ICU, and also for analysis of glycaemic state in infants at risk of hypoglycaemia in the first two days of birth. The first part of this thesis seeks to evaluate and model the different types of errors that can be included in CGM sensor measurements. The sensor model is then used to simulate the sensor in a virtual GC environment, to evaluate the trade-offs between safety, performance and workload for clinicians in the ICU when using CGM sensors to guide GC, compared to traditional intermittent POC measurements.

The second part of this thesis aims to evaluate the current metrics used to evaluate glycaemic variability, and presents a novel method for quantifying state and variability from CGM sensor traces. It presents the test cases for a cohort of new born infants at risk of hypoglycaemia, in the first 48 hours of their birth. The method is then tested using ROC curves to determine if the changes in state or variability experienced by the infants, as measured by the new method, were correlated with clinical outcomes.

Chapter 2 reviews the current metrics used to characterise and measure glycaemic variability, and presents the advantages and disadvantages of using each different metric.

Chapter 3 presents the background of CGM sensors that have been developed recently, particularly in the last decade, and assesses their use in clinical settings to date. The potential for the incorporation of a CGM device into the STAR GC protocol is proposed, and alternative control approaches are also presented.

Chapter 4 develops an autoregressive model for a CGM using two orders, one for the half hourly drift experienced by most CGM sensors, and another to model the sensor fluctuations that are present in CGM sensor readings. These models, as well as the probability noise models for both sensor drift and sensor flux, are presented.

Chapter 5 presents the results from simulating a CGM sensor and using the simulated measurements to guide GC in a virtual environment. The trade-off between GC safety, performance and workload are assessed against results from traditional intermittent POC guided GC.

Chapter 6 develops a method for characterising glycaemic variability from CGM sensor traces. The method is applied to a cohort of infants being glycaemically monitored in the first 48 hours since birth. Glycaemic States are characterised for each infant retrospectively, and examples of the method and characterisation are presented.

Chapter 7 examines the correlation between Glycaemic State Changes and negative neurologic development outcomes for the same infant cohort. Receiver operating characteristic (ROC) curves are

generated for the cohort to determine if there is any correlation between the number of State Changes experienced in the first 48 hours since birth, and neurologic impairment at 2 years old or 4.5 years old.

Chapter 8 summarises the work carried in this thesis, and presents the major conclusions from this research.

Chapter 9 provides further potential avenues of research as extensions of this work.

Chapter 2 Glycaemic Variability

2.1 Measures of Glycaemic Variability

Though perhaps qualitatively intuitive, glycaemic variability is difficult to effectively quantify. Many reviews have tried to summarise the extensive range of glycaemic variability measures, particularly in the context of diabetes management, as well as to more fully describe their advantages and limitations [75-78]. Many metrics, and adaptations of metrics, exist, but in general most fall into the following broad categories: **1)** descriptors of middle and range; **2)** descriptors of total or summed variability or excursion length; and **3)** descriptors of time in range.

The most standard methods for describing and quantifying glycaemic variability are statistical descriptors of middle and range in data. The most common descriptor of variability is to report the standard deviation (SD) of BG measurements alongside the mean BG value [79]. Similarly, the IQR (inter-quartile range) is a non-parametric alternative, reported alongside the non-parametric median BG. These metrics are popular because of their ease of use [75]. However, the very commonly used SD is a measure of dispersion, rather than variability, and is limited in its ability to reflect the time course of BG measurements [80].

The limitations of SD for describing glycaemic variability are shown in Figure 2.1, where the mean and SD are the same for two time courses of intermittent BG measurements. In this case, the same set of BG measurements is arranged in two different patterns, one a decrease at a constant rate, and the other moves between extremes. Because both have exactly the same measurement set, both have the same mean and SD. If glycaemic variability is more truly a function of the change in BG, with more rapid changes exerting different physiological effects than slower changes, than patient outcomes could differ in these two cases. Thus, the SD is extremely limited in its ability to describe the time course of glycaemic variability, which may also play a role in the result of clinical outcomes.

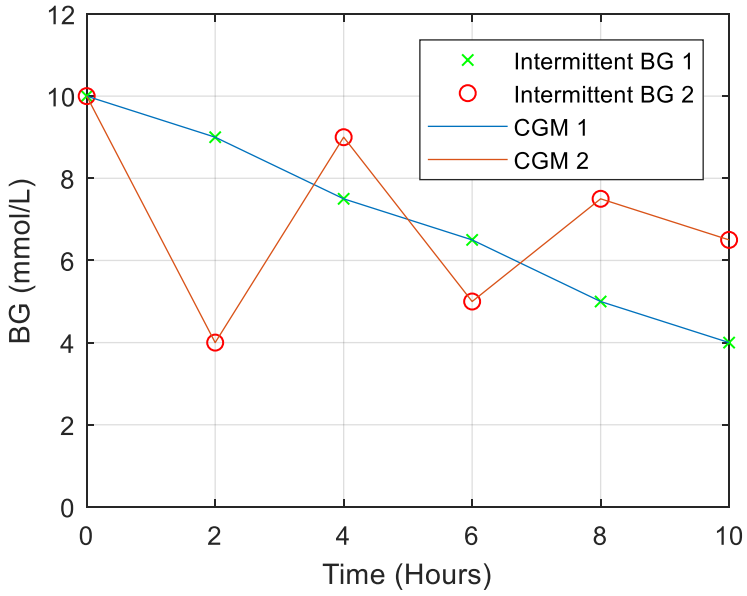


Figure 2.1: Two sensor traces with accompanying intermittent BG measures. The intermittent BG measures are the same between the two data sets, and so have the same mean and standard deviations, but display differing behaviours.

In addition, the mean and SD assume a normal distribution, which is inaccurate as BG is usually highly lognormal and skewed. Thus, this measure for variability may be inappropriate to use, particularly if there are a large number of measurements near the hypoglycaemic range of BG. Equally, all such measures requiring a normal distribution assumption have this issue, and recent reviews recommend non-parametric statistics [30].

The Coefficient of Variability (CoV), which is the mean divided by the standard deviation, is for some a preferred measure of glycaemic variability [75]. CoV can be a good measure of overall glycaemic variability, as research has shown aiming for a CoV lower than a certain threshold (36%) allows for the distinction between stable and unstable glycaemia [81, 82], where glycaemic stability is also difficult to quantify. However, this metric suffers the same limitations discussed regarding the mean and SD above. For the data in Figure 2.2, the CoV would by definition be the same between these two patients.

Metrics that look at total glucose excursion include Area under the BG curve (AUC) and Glucose Miles. AUC or area around a line is common in diabetes and non-critically ill cohort studies. It is often used with more frequently sampled continuous glucose monitoring (CGM) sensors [83-85]. Two zig-zagging sensor traces could theoretically have the same area, even though one trace is rising in level, while the other is falling. This behaviour would suggest the two traces have the same level of variability, but with differing behaviours, indicating how AUC is an overall measure of variance, but not specific to time course. This issue is also illustrated in Figure 2.2.

Glucose Miles is another way to measure variability, measuring the total 'distance' travelled by movements in BG through intermittent BG or CGM traces. It is embedded as part of some other measures [86, 87]. However, two traces can have the same Glucose Miles with very different mean BG, as shown in Figure 2.3 for a bias, and Figure 2.2 where the median is the same. Thus, this metric is similar to AUC in giving a total, but is not specific or reflective of the time-varying behaviour of blood glucose.

Furthermore, while Glucose Miles and AUC can be useful descriptors of cohort variability when paired with measurements of mean or median BG, both are unable to describe variation away from some longer term average or glycaemic state, which may be important to recognise clinically. Similar metrics, more suited to intermittent BG measurements, include the mean absolute difference (MAD), or mean of daily differences (MODD) [88, 89]. Such metrics describe the average change in BG, or the difference in BG between days, but are limited in their ability to comprehensively describe the time course of variability and its relationship to the level, as per Figure 2.3.

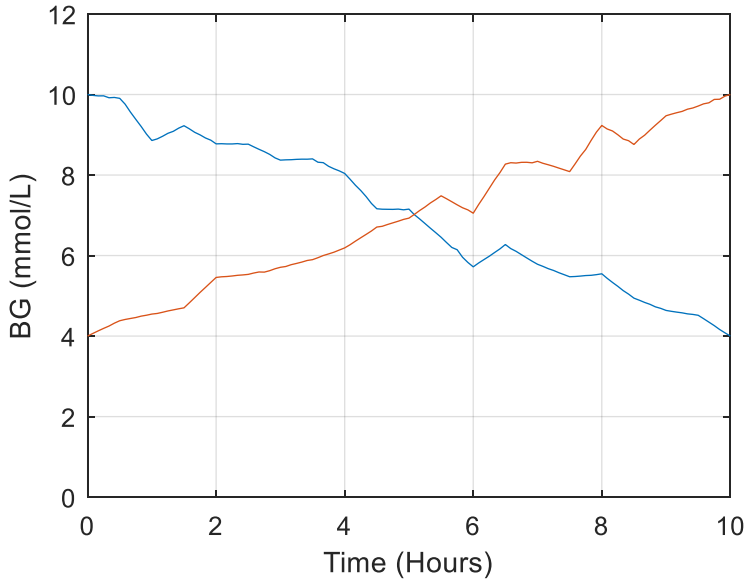


Figure 2.2: Two sensor traces with very similar AUC, but differing behaviours.

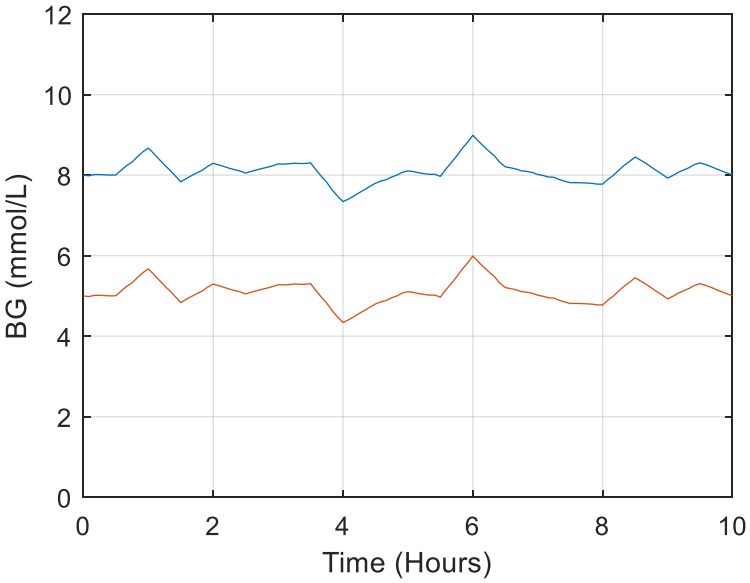


Figure 2.3: Two sensor traces with similar Glucose Miles, but with very different mean BG.

Time in band is another typical measure of variance, good for combining and capturing level and variability. However, it tends to be more of an implicit measurement of variability, rather than an explicit one. It is a good aggregate measure for larger cohorts where central tendency makes time in band informative of overall cohort behaviour. Research has shown time in band, or time in range, can be associated with clinical outcomes [25, 26, 64, 65, 90], in particular risk of microvascular

complications [91].

However, whether the measurement is inside or outside of the band itself is discrete or binary. Thus, all variability or measures in the band are assumed to be clinically acceptable, and those outside are not. Thus, a measure just inside the band is thus treated very different in analysis to one just outside, when both could be within measurement error. In addition, there is no agreement on appropriate bands, leading to difficulty in comparing the variability across studies [30]. By example, Figure 2.4 shows two sensor traces for a 4.4 mmol/L to 8.0 mmol/L range, which have the same time in band, but different, potentially clinically important behaviours both within and outside the band.

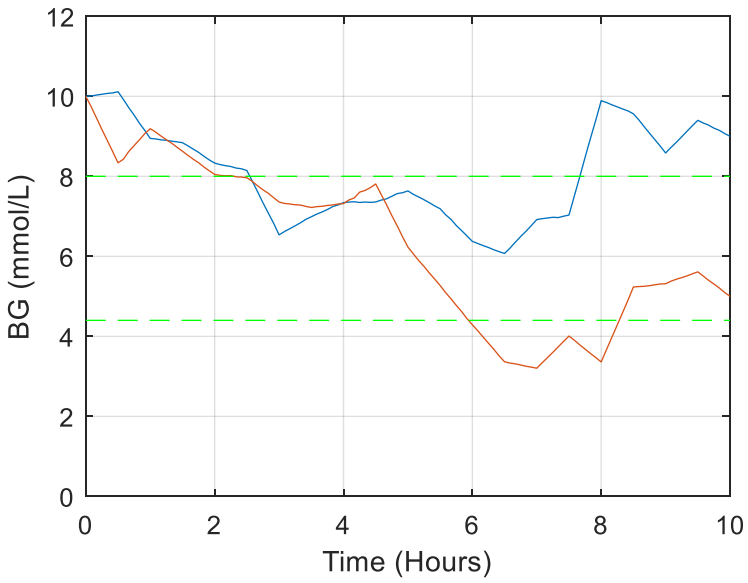


Figure 2.4: Two sensor traces with similar time in band, but differing clinical behaviour.

Other metrics for describing glycaemic variability exist, such as the mean amplitude of glucose excursions (MAGE), M-value, J-index, Low Blood Glucose Index (LBGI), High Blood Glucose Index (HBGI) and Average Daily Risk Range (ADRR) [75-78, 92]. Each of these metrics attempts to compensate for some aspect of the limitations of the metrics previously discussed, and specifically attempt to take into account aspects of desirable glucose outcomes. Most of these metrics involves one or more aspects of the middle-range, time in band, or excursion size categories. All require some degree

of correlation with outcomes to be clinically applicable, and have shown some degree of predictive capability [92]. While their disadvantages are presented elsewhere [75-78], the overall disadvantage of such metrics is that calculation and interpretation of these metrics is not transparent or intuitive to those unfamiliar with the metric, and this aspect may have impeded widespread clinical uptake.

Furthermore, most of these glycaemic variability metrics have been derived using point glucose measurements, and most have only been validated using these intermittent BG measurements to check the correlation with adverse outcomes. One of the metrics has been validated using CGM sensors, and was found to be correlated with increased hypoglycaemia given the measure of glycaemic variability over some threshold [82]. However, more studies are needed for CGM sensor derived metrics and correlation with adverse outcomes.

In addition to the various existing metrics for variability, variability in itself has been particularly hard to define in terms of what is clinically relevant. For example, observing the two sensor traces in Figure 2.1, on the timescale presented of 10 hours, the two sensor traces clearly have two differing levels of variability. Specifically, CGM sensor trace 1 is far less variable than CGM sensor trace 2. However, stretching the timescale to a 24 hour period, the two sensor traces could be interpreted in a clinical sense as having similar variability. Thus, there are currently no metrics that can adapt their timescales to take into account the full time course information CGM sensors can provide.

Table 2.1 presents a summary of all the metrics describing glycaemic variability. Their definitions and limitations are also summarised.

Table 2.1 Summary of glycaemic variability metrics, their definitions and their limitations.

Metric Name	Definition	Limitation
Standard Deviation	$\sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}}$	- Limited in its ability to describe the time course of glycaemic variability. - Assumes a normal distribution, whereas BG is normally lognormal.
Coefficient of Variability (CoV)	$\frac{SD}{\bar{X}} * 100$	- Limited in its ability to describe the time course of glycaemic variability. - Assumes a normal distribution, whereas BG is normally lognormal.
Area under Curve	Integrated area under CGM sensor trace.	An overall measure of variance, but not specific to time course as two traces could have the same area, but different behaviours.
Glucose Miles	Total glucose excursion by the CGM sensor trace.	An overall measure of variance, but not specific to time course as two traces could have the same distance travelled, but different behaviours.
Mean of Daily Differences (MODD), MAD and similar.	Describe the average change in BG or difference in BG between days.	Limited in its ability to describe the time course of glycaemic variability and its relationship with level.
Time in Band (TiB)	Calculate the percentage of glucose measurements within a desired band.	- No agreement on appropriate band. - Measurements inside or outside band are binary.

2.2 Summary

In summary, the large range of glucose metric definitions are confusing, contentious, and complicated. No metric adequately describes the time-course of glycaemia. Variability in particular is poorly captured by commonly used metrics, most of which are mathematically, rather than clinically, centred and defined. These limitations are particularly important in the context of emerging CGM technologies, which provide greater time resolution with increasingly enhanced point accuracy [93]. CGMs thus offer the potential to significantly improve GC. However, to obtain this benefit, the best clinically defined metrics to capture patient glycaemic level and state must be established, which may require a new approach to defining variability that is clinically focused rather than mathematically focused.

Chapter 3 CGM Technology

3.1 CGM Device Background

CGM devices show significant as yet untapped potential for improving glycaemic monitoring and control in the ICU, particularly due to their high measurement frequency. CGM sensors measure BG near continuously, with new generation devices able to measure at a rate of 1-6 times per minute [93-96], and standard devices providing measurements every 5 minutes [97-100]. Compared to 1-6 hourly point-of-care (POC) measurements in a well-staffed ICU [101], these devices offer huge potential to improve care and reduce workload.

The increased measurement frequency has many benefits for care, including the ability to monitor patient condition and, importantly, the trajectory of their condition in real time. They also provide warning for hypoglycaemic events [87, 102], allowing early correction. Both benefits cannot be achieved with intermittent BG measures, where the minimum clinically feasible regular measurement interval is ~1 hour. However, clinical non-compliance can be high even at lesser rates [103-105], as seen in a recent study where protocolised measurement interval was 1 hour and the clinical measurement was closer to 3 hours [31, 39].

Even at a 1-hour interval, a patient's condition may change significantly between hourly POC BG measures. These rapid changes can result in hypoglycaemia remaining untreated for 50 minutes or longer. Equally, sudden changes resulting in hyperglycaemia, indicating change in patient state, can be missed as well [106]. A CGM sensor, on the other hand, can alarm both at occurrence and predictively before it occurs, reducing risk, and potentially improving and personalising care.

Despite these advantages, CGM sensors are not widely used in the ICU as CGM sensor technology still suffers limitations, including larger point error inaccuracies and sensor drift [107-114]. They are also

expensive, so it may be hard to justify the cost versus potential benefits. The larger point accuracy errors over traditional intermittent BG measurement techniques have been well documented, with new CGM devices usually reporting gradually improving MARD values down to 8-12% [93, 115-117] as the technology has developed [93, 95, 98, 109, 115-125]. These values still do not match the 5% or lower errors common in intermittent POC measures [126]. Sensor drift is also still not widely recognised, even though it has shown to be a key driver of larger MARD and potential hypoglycaemia when used in GC [107, 127].

CGM technology also has the potential to reduce GC related workload in the ICU, providing more bedside data with lower blood sampling requirements [127-130]. Clinical practices have cited high workload as a reason for reduced intermittent BG measures, as not every ICU has a low enough patient to staff ratio to justify the increased workload some GC protocols may require [68, 70, 71]. CGM sensors are able to give a continuous readout of the current BG measurement, and clinicians can then adjust nutrition and insulin accordingly. Reduced workload and risk may also make GC more ergonomic and feasible in terms of direct and indirect burden on clinical staff [103]. However, this benefit has not yet reached regular care due to the limitations.

3.2 CGM Devices and their Use to Date

Various types of continuous glucose monitoring (CGM) device [100, 121, 131-133] have been developed over the last decade to continuously monitor BG and reduce measurement and sampling frequency. Further research into CGM devices has led to some sensors being implemented in clinical trials [79, 117, 123]. However, some recently developed CGM devices still have significant limitations in terms of accuracy and reliability [95], limiting uptake as a standard of care.

Multiple types of CGM devices have been developed differing in the site and science used for BG monitoring. Minimally invasive subcutaneous CGM devices have been developed primarily for use in type I diabetes subjects [134, 135]. These CGM devices have also been trialled in critically ill patients as substitutes or complementary sensors for intermittent BG measurements in intensive insulin therapy, reporting positive reductions in key areas of GC such as hypoglycaemic events and nurse workload [129, 130, 136-139].

In particular, Boom et al. conducted a randomized controlled trial utilising the Freestyle Navigator™ (Abbott Diabetes Care, Alameda, CA, USA) to guide glycaemic control in an ICU [130]. The Freestyle Navigator™ is one device initially designed for use in type I diabetes. Boom et al. concluded the use of a subcutaneous CGM device to guide insulin treatment was as safe and effective as intermittent measurement, and suggested designing GC protocols to work in conjunction with CGM may increase GC performance.

Holzinger et al. conducted a randomised controlled trial testing the CGMS® System Gold™ (Medtronic, Northridge, CA) in an ICU setting [140]. The glycaemic control performance and safety of CGM devices were not directly evaluated in this study. Instead, the main aim was to evaluate the impact of shock requiring norepinephrine on the accuracy and reliability of continuous glucose monitoring. Again, this CGM device was initially designed for use in type I diabetes, but the limitations of using subcutaneous CGM devices in an ICU setting were explored in this study. They concluded circulatory shock requiring norepinephrine therapy had no influence on the accuracy and reliability of the CGM device in critically ill patients.

Wollersheim et al. undertook a prospective clinical trial with the Medtronic Sentrino® CGM system, a CGM device specifically designed for use in ICU patients [95]. This CGM device showed accuracy in the continuous data display, with a mean absolute relative difference (MARD) of < 14%. However, due

to factors such as low point accuracy and prolonged data gaps, the CGM device was reported to not perform with satisfactory accuracy, feasibility and clinical compliance.

One of the potential limitations of these CGM devices is they measure BG indirectly from interstitial glucose. A physiological time lag exists between the BG concentration and the interstitial glucose concentration due to diffusion through the capillary wall. However, studies have shown the maximum average time lag is around 10 minutes [141, 142]. Thus, for the purposes of GC, where the time between interventions may range from 1 to 3 hours, the physiological time lag between interstitial glucose and BG is not likely to be a source of significant error.

Other types of more invasive intravascular CGM devices have been developed to measure BG directly via an arterial line [143-145]. Although these CGM devices are more invasive, many critically ill patients may already require an arterial line for other treatments. Thus, measuring BG more directly through this existing arterial line would provide the added benefit of more accurate BG measurements required for GC with no additional invasiveness. These CGM devices are not yet widely used or validated in ICUs, or other patients.

3.3 In-Silico CGM Protocol Testing at the University of Canterbury

A GC protocol testing framework has also been developed at the University of Canterbury bioengineering department to evaluate the performance and safety of GC protocols before they are implemented in a clinical setting. Virtual patients are used in lieu of real critically patients, which are generated using STAR clinical data from Christchurch Hospital ICU. Intermittent BG and insulin-nutrition interventions are combined with the physiological model to fit an hourly insulin sensitivity (SI) value [146]. The result is an SI profile over time for the duration of the patient's glycaemic control, which reflects their underlying metabolic responsiveness [49, 53, 147, 148]. As SI is treatment

independent [149-151], it can be used to simulate the results of other interventions in-silico. A virtual patient is thus defined by a time ranging SI profile developed from clinical data [149, 151]. This virtual trial method has been clinically validated on independent data, in clinical use, and in prediction of trial outcomes across multiple cohorts [62, 149, 150, 152].

Within this framework, there is potential to develop a virtual CGM device model for use in virtual trials to replace the virtual intermittent BG measurements currently utilised for the STAR protocol. A similar study has developed a CGM device model that simulates the measurement of BG via interstitial glucose of a single sensor, the FreeStyle Navigator® CGM [153]. The model incorporated diffusion of BG from the blood plasma to the interstitial fluid, as well as additive autoregressive moving average noise. The simulated sensor was built to use with a virtual GC protocol testing framework built by the University of Virginia. However, to date no such CGM device modelling work has been undertaken to incorporate the virtual CGM device with STAR or any established glycaemic control framework.

3.4 Alternative CGM-based Control Approaches

Other types of control also exist that may benefit from the enhanced temporal resolution a CGM sensor would provide. There are three broad categories of control, control that is reactive and looks at previous measurements to control level, control that is predictive and looks forward to control level, and control that is also predictive and looks forward, but doses on risk to directly manage variability more than level [30].

Reactive PID control has been trialled for GC in the past. Chee et al. [154, 155] conducted initial trials in a small pilot study and were not able to adequately lower glucose levels with their PID controller. Some of the issues encountered included sensor failure and large sensor errors. With recent

developments in CGM technology, the sensor error and failure rates have improved markedly, but accuracy may not yet be adequate for a PID control framework.

A model predictive controller, such as the one used by Plank et al. [56], uses a physiological metabolic model to predict the future level of BG given an insulin-nutrition intervention. The benefit of this prediction is the glycaemic result can be determined for a range of insulin and nutrition levels, and thus the most appropriate intervention chosen based on goal feed levels and glucose band targets. However, this controller targets BG level and does not account for variability.

The STAR protocol builds upon the model predictive controller by predicting future insulin sensitivity and its likely changes [51-53, 156, 157]. As a result, future BG and likely variability in BG outcomes is directly quantified, allowing the risk of hypo- and hyper-glycaemia to be directly managed in GC [34, 54, 61, 62, 156, 158]. Managing variability, and thus improving control, can improve mortality rates in the critically ill as variability is directly associated with outcome [21, 22]. The stochastic model built into the STAR protocol directly addresses the issue of patient variability by dosing on risk of hypoglycaemia to improve control and glycaemic outcomes.

3.5 Summary

Overall, CGM technology has not reached regular clinical use in ICU care, despite common acceptance in lower-risk outpatient type 1 diabetes care [159-163]. As CGM technology improves it will be able to replace intermittent measures used in the ICU today. However, with increased temporal measures and benefits to control, the same issues of how to quantify level and state remain. In fact, with the increased measurement rate of CGM technology, these issues are exacerbated as measures that work for intermittent measures may not be representative or accurate for use with CGMs. There thus remains a significant need for better, more clinically defined metrics, and in particular, those able to maximise the

measurements delivered by CGMs.

Chapter 4 CGM Modelling

4.1 Background

CGM sensor technology has potential to improve GC, but the increased temporal measurement resolution CGM devices provide is still, at least somewhat, outweighed by the larger point accuracy errors in these devices due to sensor drift, bias, and random noise [108, 113, 128, 153, 164-166]. Trend accuracy is also an important factor, particularly where alarms indicating hypo-, and hyper-, glycaemia are concerned [167]. As a result, GC utilising CGM in the ICU is very uncommon, and is not yet standard practice anywhere, despite some studies showing it may be feasible [129, 130, 168-172].

There is thus a need to model and account for sensor error, preferably in a generalisable modelling method. Such a method would in turn enable optimal (model-based) design of GC protocols maximising CGM advantages and minimising their disadvantages in GC. It would also match recent consensus statements from medical and industry based working groups that measures of accuracy need to be standardised for licensing authorities, and to enable comparisons across studies and devices [173, 174].

Three CGM models have been developed in the past, primarily for interstitial CGM devices, by Breton and Kovatchev [153], Lunn et al. [113], and Facchinetti et al. [108]. The model developed by Breton and Kovatchev has been used in *in-silico* preclinical trials to simulate the effectiveness of using an interstitial CGM device for closed-loop control in type 1 diabetes [175, 176]. However, Facchinetti et al. have since shown the modelling methodology used by Breton and Kovatchev may be sensitive to small errors in CGM data calibration or errors in the description of BG-to-interstitial glucose (IG) dynamics [110]. Lunn et al. [113] developed a more refined version of the model developed by Breton and Kovatchev by fitting a dynamic model with forcing functions. However, it suffers from the same issue as Breton and Kovatchev's model. In addition, neither model separately considered sensor drift,

thus including it in point error. They were thus not as accurate for sensors where drift occurs, which is all current approved (chemistry-based) CGM sensors.

Facchinetti et al. further developed a model of sensor error incorporating BG-to-IG dynamics, using an autoregressive model to account for additive measurement noise and a linear time-varying model to account for calibration and sensor drift [108]. Again, BG-to-IG dynamics are added because the CGM device modelled was an interstitial device. This modelling method was able to account for sensor drift, and also split sensor error into multiple components, including error arising from calibration and measurement noise. However, this modelling method is very data and labour intensive, requiring multiple CGM devices per patient and 15 minute intermittent BG measurement intervals. This large amount of data may not be available in past data acquired from CGM device trials, while the large workload required to measure BG every 15 minutes would be a clinical and feasibility barrier for further sensor error characterisation [70, 71, 104, 105], and thus to modelling and simulation of sensor behaviour in new CGM devices.

Additionally, none of the previous modelling efforts have taken into account trend accuracy of the CGM devices studied. Signal et al. [167] developed the Trend Compass and Trend Index to assess a CGM sensor's trend accuracy. In particular, trend accuracy could be seen as equally or more important as the measurement of mean absolute relative difference (MARD) to assess point accuracy in the previous modelling efforts, especially because trends are often used in clinical decision making.

The GlySure (GlySure Limited, UK) CGM device considered in this work is from a newer class of ICU devices measuring venous BG via an intravenous line, thus avoiding IG dynamics. The sensor is comprised of microporous and dialysis membrane, hydrogel, optical fibre and a thermocouple, while the glucose detecting chemistry used is a fluorescent diboronic acid receptor, embedded within the

hydrogel. Placement of the sensor itself can occur through either a central venous catheter, or a radial artery catheter [93].

This chapter presents a novel auto-regressive (AR) method and model characterising this CGM sensor. The modelling method is capable of characterising the CGM sensor with less data than required by previous characterisation methods in terms of sensors per patient, and explicitly accounts for sensor drift, while maintaining both the point and trend accuracy of simulated sensor errors. This model is developed and compared to clinical data to assess its validity. The overall modelling method and approach is generalisable to similar devices.

4.2 Methods

4.2.1 Sensor Modelling

4.2.1.1 Clinical Data

Data was sourced from an observational pilot trial of the CGM device on 33 cardiac intensive care patients (duration 21-51 hours per patient), where CGM readings were not used clinically for GC. The sensor provides a new reading 4 times per minute, and intermittent BG measures are used to calibrate the sensor (calibration or recalibration BG) approximately every 8 hours. Intermittent independent BG measures not used to calibrate the sensor (reference BG) were taken approximately every 2.5 hours. Each intermittent BG measurement was taken using YSI 2300 STAT Plus (Yellow Springs Instruments, Yellow Springs, OH) or the i-STAT (Abbott Laboratories, Abbott Park, IL), which are highly accurate measures, in order to minimise the error in reference BG values (Mean Percentage Error of 1.79 %) [177, 178], as reported in [93].

Patient details and details of the pilot trial can be found in Table 4.1 and Table 4.2. Data from Patients 10 and 22 were later discarded due to sensor failure. The mean absolute relative difference (MARD) was calculated between paired BG (calibration and reference) and CGM measurements. The average global MARD (excluding Patients 10 and 22) was 9.6%.

Table 4.1: Patient details from the Glysure CGM sensor clinical trial. The mean (range) is reported for the Duration, BMI and Age.

Characteristic	Cardiac patients
n	33
Duration (hours)	40.8 (21.1-50.7)
Male	22 (66.7%)
Female	11 (33.3%)
Individuals with Diagnosed Diabetes	14 (42.4%)
Hypertensive	15 (45.5%)
BMI	25.3 (17.7-35.8)
Age	50.8 (19-77)

4.2.1.2 Sensor Characterisation

Sensor characterisation uses two independently defined AR models to separately capture drift and higher frequency sensor fluctuations, where most other methods have not explicitly accounted for sensor drift [113, 153]. This method is able to be used on the clinical data where there was only one CGM sensor per patient, which would not have been possible with the method of [108], as the characterisation of separate noise components requires multiple sensors for each patient.

Table 4.2: Pilot trial details.

Patient Number	No. Hours of CGM	No. Calibration BG	No. Reference BG	MARD	Median [IQR] Sensor Error (%)
1	48.33	3	14	8.62	8.4 [4.6 12.3]
2	40.97	3	13	6.20	5.1 [1.7 10.4]
3	42.66	3	13	6.00	2.9 [1.7 9.4]
4	47.33	2	15	10.66	9.7 [6.3 14.3]
5	47.05	3	15	6.85	4.6 [2.0 11.2]
6	50.69	3	14	9.57	9.3 [7.8 10.8]
7	46.90	3	14	12.56	13.3 [6.5 16.8]
8	43.43	3	14	4.47	2.7 [1.7 5.9]
9	47.05	7	14	13.25	10.9 [3.5 18.2]
10	-	-	-	-	-
11	48.69	3	15	15.62	8.5 [4.1 11.2]
12	42.84	3	13	11.73	6.1 [1.5 12.0]
13	40.91	3	12	6.10	14.5 [10.8 25.4]
14	45.60	4	13	12.10	13.0 [8.1 21.3]
15	39.37	5	10	4.60	12.8 [6.0 18.3]
16	43.51	7	9	6.43	6.9 [2.7 13.7]
17	37.71	3	12	7.48	10.3 [6.1 21.5]
18	36.72	5	11	7.84	11.3 [9.1 15.7]
19	39.23	2	13	4.85	4.5 [2.7 8.5]
20	40.42	4	11	7.19	4.0 [2.3 7.2]
21	36.35	4	11	8.27	7.2 [4.0 7.9]
22	-	-	-	-	-
23	37.77	3	13	5.29	6.0 [2.7 10.3]
24	39.03	4	12	12.20	3.5 [1.5 8.0]
25	38.12	3	13	4.77	12.8 [5.0 17.4]
26	36.69	3	12	6.22	4.4 [2.2 6.4]
27	40.12	4	12	12.36	4.7 [3.2 7.7]
28	37.65	4	12	14.49	8.5 [7.7 16.0]
29	37.81	3	12	22.93	16.5 [9.7 20.9]
30	36.51	4	11	10.12	21.9 [18.5 27.7]
31	37.54	4	12	8.19	8.3 [3.3 13.9]
32	21.14	2	7	17.38	7.0 [3.4 11.8]
33	25.49	3	9	14.35	14.0 [11.9 17.7]
Overall Cohort Median and Range	40.12 (21.14 - 50.69)	3 (2 - 7)	12 (7 - 15)	8.27 (4.47 - 22.93)	-

*Patients 10 and 22 were discarded in this retrospective analysis due to sensor failure

A. Drift:

Clinical data is divided into separate periods between recalibration points for each patient. Drift is characterised for any given patient trace between recalibration measurements using the percentage difference between sensor and reference measurements, as assessed half-hourly using sensor glucose (SG) ($BG_{SG/30\,min}$), and intermittent BG interpolated ($BG_{IM/30\,min}$) between calibration and reference

measurements (BG_{IM}). Interpolated intermittent BG_{IM} measures are a vector with an interpolated value every 30 minutes, defined:

$$BG_{IM/30\,min} = interp(BG_{IM})|_{t=0:30:t_{end}} \quad (4.1)$$

Half hourly sampling of the CGM sensor trace simply takes the paired CGM value at that time:

$$BG_{SG/30\,min} = sample(BG_{sensor})|_{t=0:30:t_{end}} \quad (4.2)$$

The half hourly interval matches observed physiological and clinical time frames for BG fluctuations and trends, and thus captures the broad trends caused by sensor drift and, over several samples, eliminates the impact of random errors at any given point. Thus, a potential percentage drift can be calculated at each paired half hourly value:

$$Drift = \frac{BG_{SG/30\,min} - BG_{IM/30\,min}}{BG_{IM/30\,min}} \quad (4.3)$$

A lag-2 AR model is then used to characterise the observed drift for a given patient's data or for the whole cohort. This AR model uses the entire cohort's data (N = 31 patients), and is defined:

$$Drift_{n+1} = \alpha_d + \beta_d * Drift_{n-1} + \gamma_d * Drift_n + \xi_d \quad (4.4)$$

Model parameters α_d, β_d and γ_d are identified using linear least squares from the half hourly [$Drift_{n+1}$, $Drift_n$, and $Drift_{n-1}$] data points derived from the entire clinical data cohort (N = 31 patients) and Equations 4.1-4.3, assuming ξ_d (or drift Eta) = 0. The parameters α_d, β_d and γ_d thus capture sensor drift behaviour over the entire cohort.

Having identified the best fit (α_d, β_d and γ_d) for the cohort, a drift noise term is calculated by re-arranging Equation 4.4 and solving for the residuals, ξ_d. Outlier drifts from Equation 4.3 of more than ±25%, where 99.3% of the data is within this threshold, were discarded to maintain sufficient data density for probability modelling. These results are used to create a drift noise model by smoothing a continuous distribution function across the ξ_d residual results obtained by rearranging Equation 4.4.

B. Random Sensor Noise:

Random sensor noise fluctuations are assessed at the CGM sensor sample rate (4 measures per minute) and are sampled from the interpolated BG and CGM device measurements every 15 seconds or 0.25 minutes, yielding:

$$BG_{base/0.25\,minutes} = interp(BG_{SG/30min})\big|_{t=0:0.25:t_{end}} \tag{4.5}$$

The fractional difference between these linearly interpolated 'base' BG points and the real sensor trace is defined:

$$SensorFlux = \frac{BG_{sensor} - BG_{base/0.25\,minutes}}{BG_{base/0.25\,minutes}} \tag{4.6}$$

Another lag-2 AR model is used to characterise these sensor fluctuations around the trend, defined:

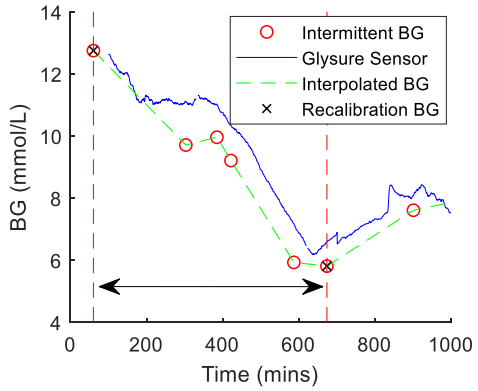
$$SensorFlux_{n+1} = \alpha_{sf} + \beta_{sf} * SensorFlux_{n-1} + \gamma_{sf} * SensorFlux_n + \xi_{sf} \tag{4.7}$$

The model parameters α_{sf}, β_{sf} and γ_{sf} are identified using the entire data cohort (N = 31 patients) and linear least squares from [$SensorFlux_{n+1}$, $SensorFlux_n$, and $SensorFlux_{n-1}$] data points obtained every 15 seconds (0.25 minutes), as derived from clinical data, and Equation 4.6, assuming ξ_{sf} (or sensor fluctuations Eta) = 0.

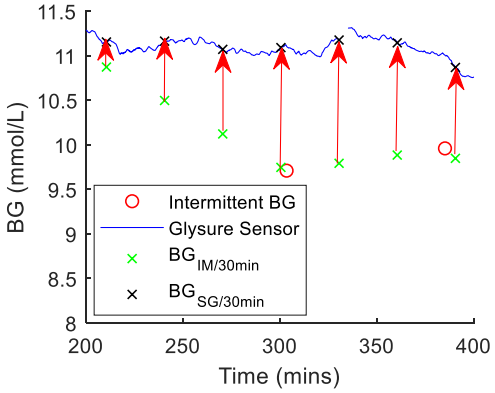
After identifying a linear best fit (α_{sf}, β_{sf} and γ_{sf}) across the entire data cohort, the sensor fluctuation noise term is calculated by re-arranging Equation 4.7 and solving for the residuals, ξ_{sf}. Outlier fluctuations of more than ±1%, where 99.9% of the data is within this threshold, were discarded. These results are used to create a sensor fluctuation noise model, similar to the drift noise model, by smoothing a continuous distribution function across the data range of ξ_{sf}.

C. Illustrated Example of Sensor Characterisation:

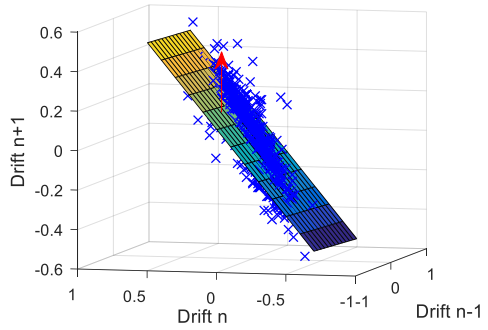
An example of the modelling process is shown using data from Patient 2 in Figure 4.1. Figure 4.1(a) shows the clinical sensor data for Patient 2 over the first 16 hours, with the data split into separate periods between recalibration times to characterise drift. Figures 4.1(b) and 4.1(c) show characterisation of the AR drift model using Equations 4.3 and 4.4, where Figure 4.1(c) shows the plane of best fit for model parameters α_d, β_d and γ_d, and the drift noise term definition (ξ_d). Figures 4.1(d)-4.1(f) show characterisation of the sensor fluctuation model using Equations 4.6 and 4.7, with planes of best fit for model parameters α_{sf}, β_{sf} and γ_{sf} and the sensor fluctuations noise term definition (ξ_{sf}).



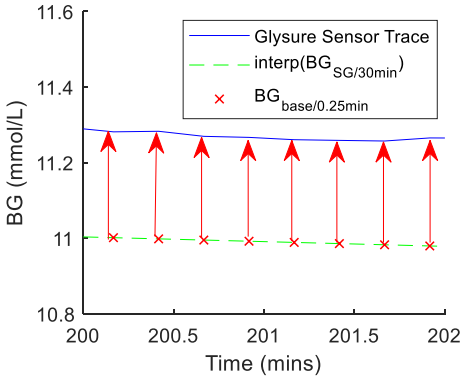
a) Clinical data plotted for Patient 2.



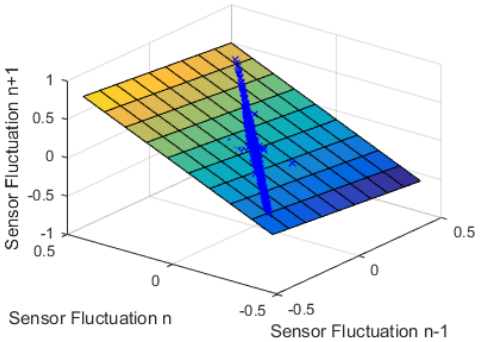
b) Patient 2 with BG resampled every half hour. Red arrows show how drift was characterised (Drift = $(BG_{SG/30min}-BG_{IM/30min})/BG_{IM/30min}$)



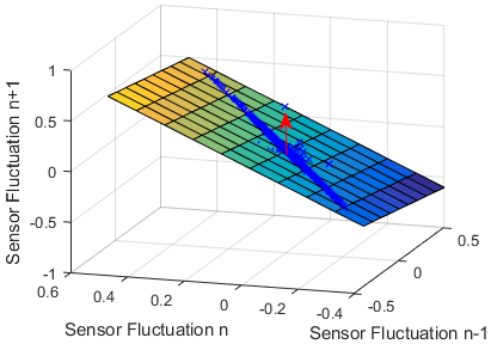
c) $Drift_{n+1}$ plotted against current ($Drift_n$) and previous ($Drift_{n-1}$) for all patients. The plane of best fit is shown in black, with a red arrow showing how the residual ξ_d is sampled.



d) Patient 2 with BG resampled every minute. Red arrows show how the sensor fluctuation was characterised (Note small x and y scales).



e) Sensor $fluctuation_{n+1}$ plotted against current sensor fluctuation (n) and previous sensor fluctuation (n-1) for all patients. The plane of best fit is shown.



f) Sensor $fluctuation_{n+1}$ plotted against current sensor fluctuation (n) and previous sensor fluctuation (n-1), with a red arrow showing how the ξ_{sf} is sampled for the fluctuations.

Figure 4.1: Steps of the sensor characterisation process, showing how the drift and sensor fluctuations were characterised.

4.2.2 Sensor Simulation

Sensor simulation for validation simulates a sensor trace given intermittent BG_{IM} measures in a process essentially the reverse of sensor characterisation. Intermittent BG_{IM} measures are used as a base to simulate the CGM sensor, as that is all the clinical data that might be available when simulating a virtual patient to be monitored by CGM [149, 179-181]. They are interpolated half hourly, and drift applied using Equation 4.4. The resulting BG with drift is defined:

$$BG_{base(t=0:30:t_{end})} = BG_{IM/30\ minutes} * (1 + Drift) \tag{4.8}$$

Where *Drift* is calculated using Equation 4.4, with ξ_d drawn randomly from the cohort probability distribution generated from the clinical data. At t = 0 and any calibration BG, approximately every 8 hours for this device, the condition [$Drift_n$, $Drift_{n-1}$] = [0.0, 0.0] is used to recalibrate the CGM sensor to match the simulated intermittent BG measurement, providing a point to point calibration. This approach could be modified to account for any more complex calibration process.

BG_{base} is then linearly interpolated to provide enough data points to match this CGM device's 4 times per minute rate, such that:

$$BG_{base/0.25\ minutes} = interp(BG_{base})|_{t=0:0.25:t_{end}} \tag{4.9}$$

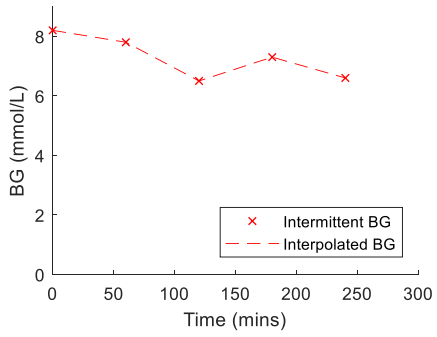
The simulated sensor output is then defined:

$$BG_{new\_sensor(t=0:0.25:t_{end})} = BG_{base/0.25\ minutes} * (1 + SensorFlux) \tag{4.10}$$

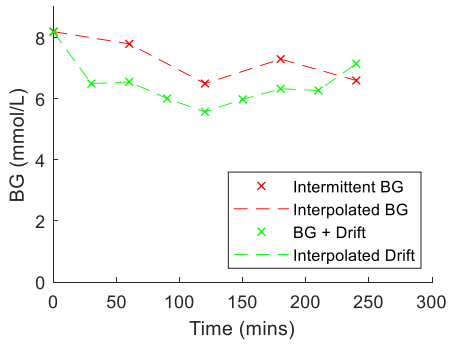
Where *SensorFlux* is calculated according to Equation 4.7, with ξ_{sf} drawn randomly from the cohort probability distribution generated from clinical data.

Once again, at t = 0 and any calibration BG, the condition $[SensorFlux_n, SensorFlux_{n-1}] = [0.0, 0.0]$ is used to recalibrate the CGM sensor. Finally, an additional limit of a maximum drift of 40% was applied to the drift AR model matching extremes seen in the clinical sensor data.

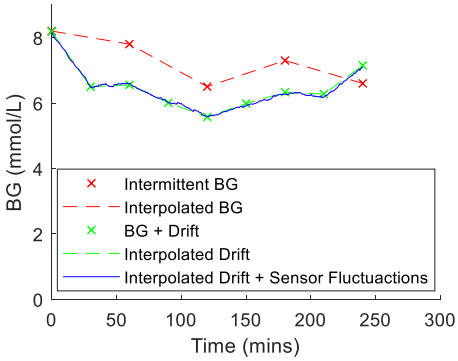
Regarding calibration, setting $[Drift_n, Drift_{n-1}] = [0.0, 0.0]$ and $[SensorFlux_n, SensorFlux_{n-1}] = [0.0, 0.0]$ effectively applies a point to point recalibration, allowing the sensor trace to go through a recalibration point. Divergence from the interpolated BG is then initiated by the ξ_d and ξ_{sf} terms. Figure 4.2 outlines the steps during simulation, with a sensor trace and virtual patient forward simulated for the first 4 hours given the initial intermittent BG data in Figure 4.2(a).



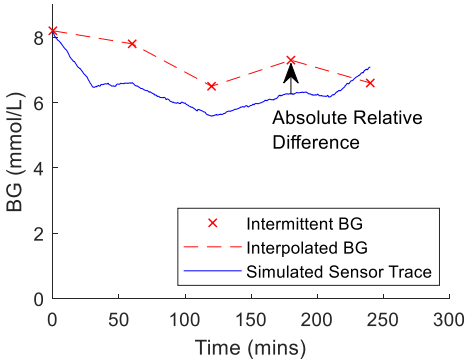
a) An example patient. Intermittent clinical BG measures are interpolated.



b) AR drift is applied to the interpolated BG at half hour intervals (Equation 4.8).



c) AR sensor fluctuations are added to the interpolated AR drift (Equations 4.9 and 4.10).



d) Comparison of intermittent BG and final sensor simulation.

Figure 4.2: Sensor simulation steps, calibration not shown.

4.2.3 Sensor Model Validation: Qualitative and Quantitative

Sensor traces are simulated using intermittent BG from the clinical data cohort from which the model was built. To test consistency between the model and the clinical sensor data, several sensor simulations were overlaid with the clinical data for each patient trace and compared in a blinded test. If the clinical sensor traces were difficult to visually distinguish from simulation, then the model was qualitatively accepted as broadly capturing key behaviour. This qualitative assessment enables assessment of trends and features not easily compared in quantitative tests.

To quantitatively validate the model, a single simulation was run to generate a single sensor trace from the sensor model for each patient to compare to the clinical sensor data. Percentage difference distributions for the simulated sensor data and the clinical sensor data were compared, and a Clarke Error Grid (CEG) plot and Bland-Altman plot constructed. The Bland-Altman plot enables analysis of any differences in bias behaviour between clinically measured and modelled sensor traces. Trend accuracy was assessed and compared for the clinical data and simulation data using the Trend Compass and Trend Index described by Signal et al. [167]. A Trend Compass and Trend Index, defined in Equation 4.11 as an absolute mean angle from perfect trend accuracy, were produced for both clinical and simulation data to validate the reproduction of trend accuracy using the simulation method.

The Trend Compass is a visual tool that can be used to quickly identify the trend accuracy of a particular CGM sensor. The compass itself is split into four separate quadrants depending on the relative changes of BG and SG, with measurements plotted according to the angle between the interpolated BG vector and the corresponding CGM vector (θ), and radially according to the last interpolated BG measurement. The angle of theta (θ) is plotted as measured from the vertical and represents the level of agreement between the rates-of-change in BG and SG, where points plotted on the vertical have perfect trend accuracy. It thus evaluates rising (upper half) and falling (lower half) trends separately, where the two sides distinguish trends where CGM measured SG rises (or falls) faster or slower than BG. Figure 4.3

illustrates this tool in full. This quantitative validation was carried out for each patient and sensor trace in the clinical cohort.

$$TI = \frac{1}{n}\sum_{i=1}^{n} |\theta(i)| \tag{4.11}$$

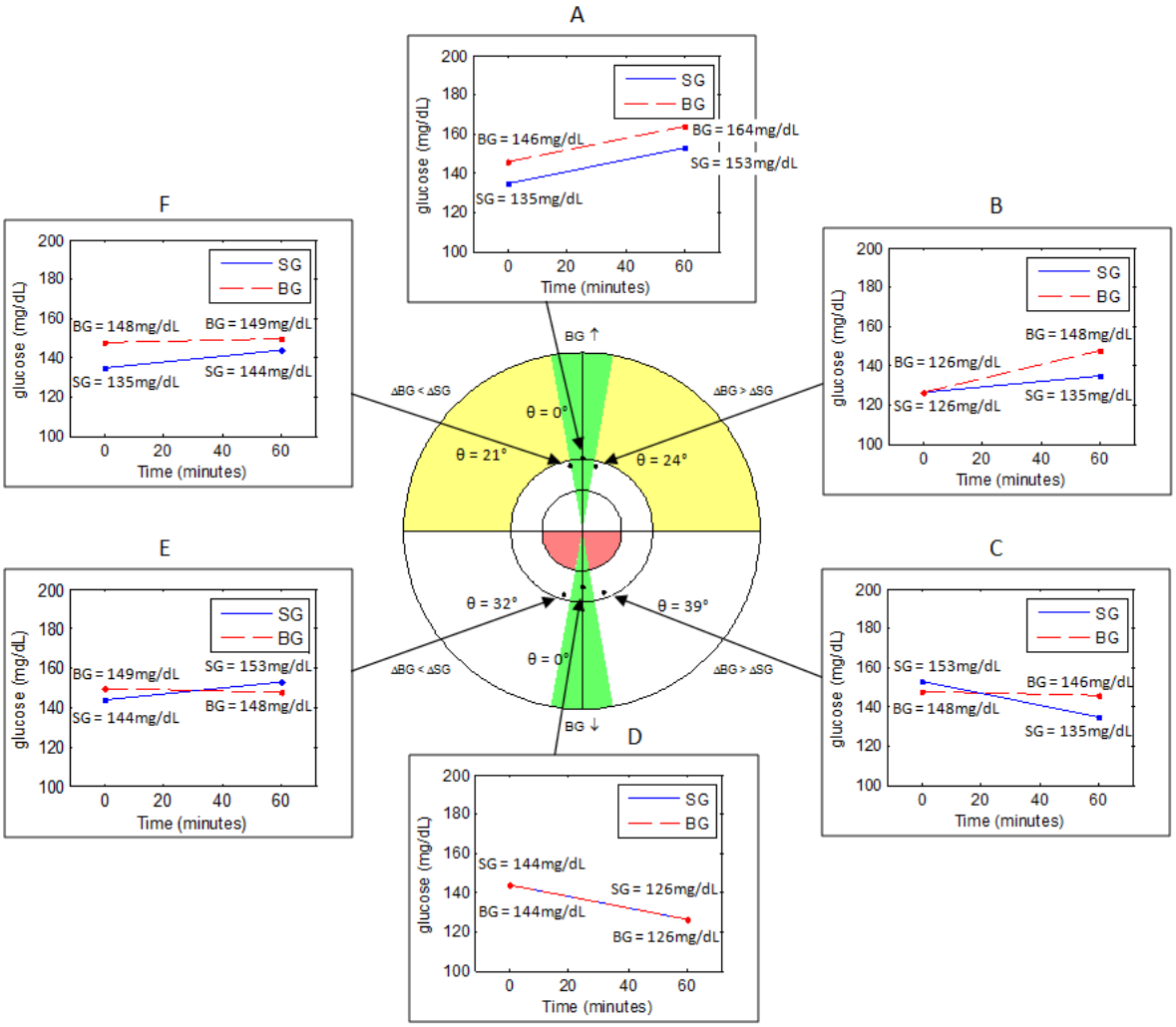


Figure 4.3: Six examples of SG and BG paired measurements with their corresponding point on the Trend Compass. Note: comparing 'A' to 'D' shows that the constant bias has no effect on how trending is displayed on the Trend Compass (Both examples have perfect trend accuracy so θ=0°).

Further Monte-Carlo simulations were undertaken, until a total of 100 simulations worth of data was generated for each patient. Individual patient simulation traces were plotted on top of the clinical data to further check the consistency between the clinical data and simulated traces. The range of simulated

sensor behaviour was also compared to the original clinical data to ensure simulated sensor behaviour reflected extremes in actual sensor behaviour. For each patient, the minimum simulated BG value was taken for each time point over the 100 simulations to generate an overall minimum sensor profile. A maximum sensor profile was also produced for each patient in a similar method, using the maximum simulated BG values for each time point. Together, these minimum and maximum profiles formed an area profile of all possible simulated BG values, which were then compared to the clinical sensor traces for each patient.

4.2.4 Convergence Analysis

To validate this model methodology is able to work with limited reference readings, the number of reference BG measurements that were not recalibration measures was halved (42% less measures), and the sensor was re-characterised for drift and noise. The global MARD for the simulations of the sensor was re-evaluated and percentage difference distribution and CEG plots generated to compare with the previous analyses.

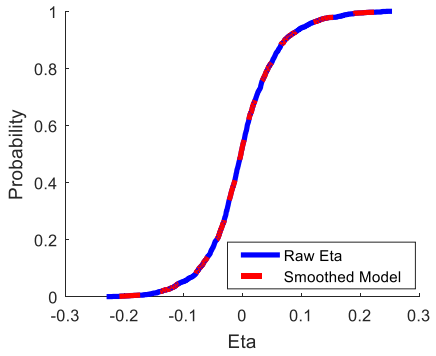
4.3 Results

4.3.1 Sensor Characterisation

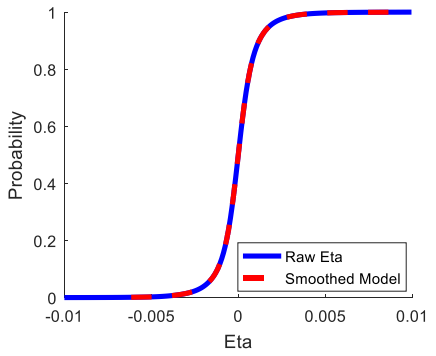
Table 4.3 gives sensor model parameters identified from the cohort data, and Figure 4.4 shows the noise term model distributions (ξ_d and ξ_{sf}), raw and smoothed fit, for drift and sensor fluctuations.

Table 4.3: Auto-regressive (AR) model parameters for drift and sensor fluctuations.

AR pass	Key characteristic	α	β	γ	ξ median	ξ max (absolute)	R^2
1	Drift	-0.00147	-0.07152	0.9698	-0.0023	0.2486	0.81
2	Sensor Fluctuations	-6.0e-6	-0.261	1.261	-2.66e-06	0.01	0.99



(a) Noise model for drift, ξ_d



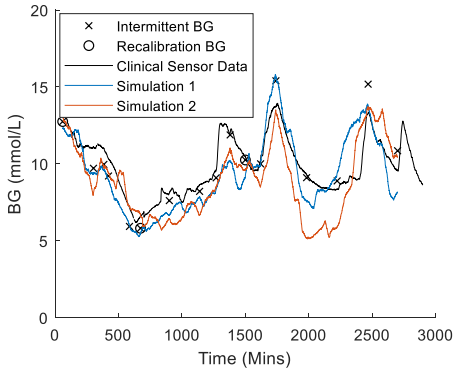
(b) Noise model for sensor fluctuations, ξ_{sf}

Figure 4.4: Random noise distributions for auto-regressive drift and sensor fluctuations (fraction).

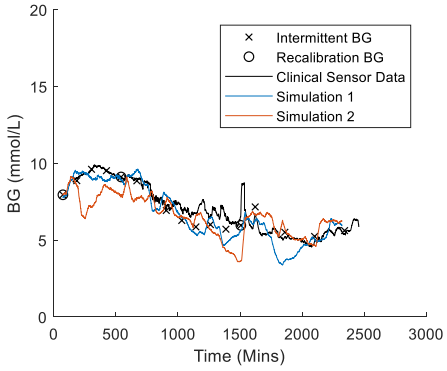
4.3.2 Sensor Simulation

Figure 4.5 shows four example patients with the real sensor trace plotted alongside two simulated traces. Sensor behaviour is visually consistent between simulated and real sensor traces, with drift and sensor fluctuations of approximately the same magnitude across all traces. In a few sensor traces, such as the clinical data in Figures 4.5(b), there is additional high frequency low amplitude noise the model cannot capture. This noise is not likely to affect glycaemic control applications of this sensor, whether in simulation or practice. The model is thus considered qualitatively good.

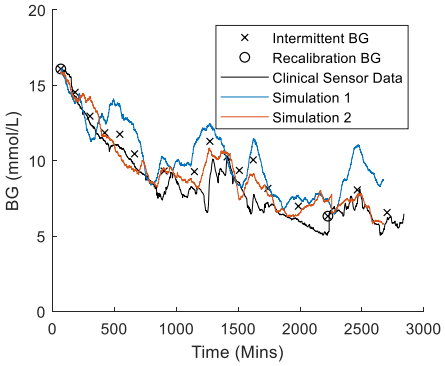
Figure 4.6 shows the distribution of percentage differences over all N = 31 patients clinical and simulation data, with one simulation per patient. The percentage differences were calculated by taking SG, subtracting the interpolated intermittent BG, and then dividing through by the interpolated intermittent BG. The model simulated data distribution is slightly tighter than the clinical data. This outcome is mainly due to the point-to-point recalibrations used in the simulations, which is a slightly more accurate recalibration than the one used in the clinical trial. There is also some slight non-Gaussian distribution of percentage differences at the extremes, compared to the Gaussian noise distribution used in the AR modelling. However, the distribution is still very similar, indicating the modelling method accurately recreates the percentage differences in the measurements of intermittent BG and SG.



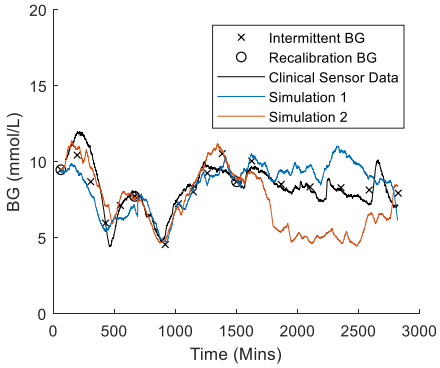
(a) Clinical sensor data plotted alongside 2 simulations for patient 1.



(b) Clinical sensor data plotted alongside 2 simulations for patient 2.



(c) Clinical sensor data plotted alongside 2 simulations for patient 4.



(d) Clinical sensor data plotted alongside 2 simulations for patient 5.

Figure 4.5: Clinical sensor data with two sensor traces generated from the model.

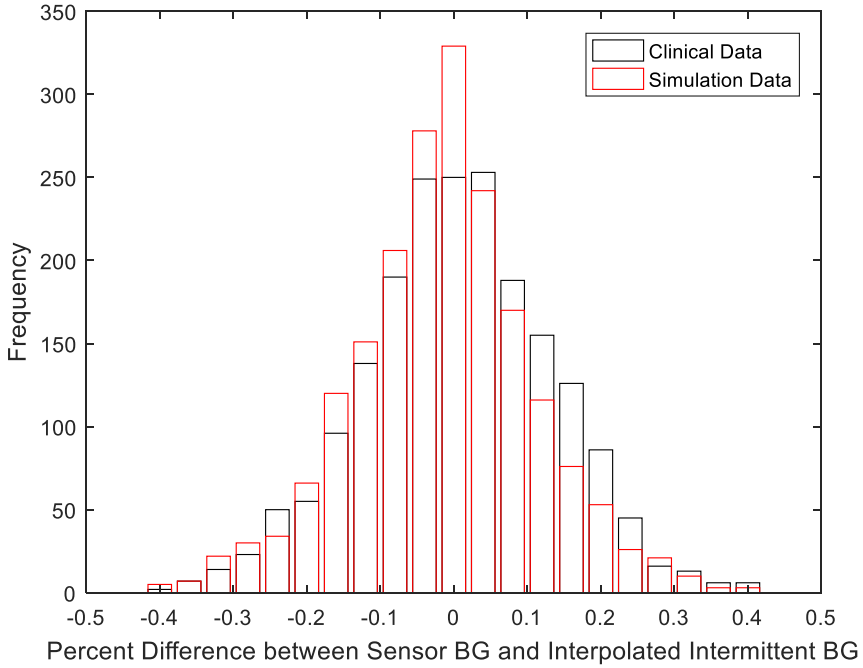


Figure 4.6: Percentage difference distribution plot of the clinical sensor data vs. the simulated data.

The Clarke Error Grid (CEG) in Figure 4.7 shows the model behaves in a consistent manner to the clinical data. While the distributions are consistent across the BG range, as expected, there are a few outliers in the clinical data not captured by the model. The percentages of measurements falling within the zones of the CEG plot are shown in Table 4.4. Slightly more data points fell within zone A than zone B when comparing the simulated results to the clinical results, suggesting the simulation method may be slightly more accurate than the clinical sensor. However, this difference is minor (~6.6% change between zones A and B), and thus the model could still be considered to quantitatively represent the sensor behaviour well. In Figure 4.8, the Bland-Altman plot shows no significant bias across the observed BG range, and the plotted lines of ±2σ for the clinical and simulated data show the strong similarity between the model outputs and the clinical sensor data over this 95% range.

Of note, there is very little clinical data below 5 mmol/L. Underlying sensor model assumptions apply constant sensor behaviour across the full BG range resulting in similar proportional BG error at high and low BG. This assumption is used for lack of other data from the sensor at this time. In this case, at lower BG, this choice translates to a consistent percentage error, resulting in slightly lower absolute BG errors.

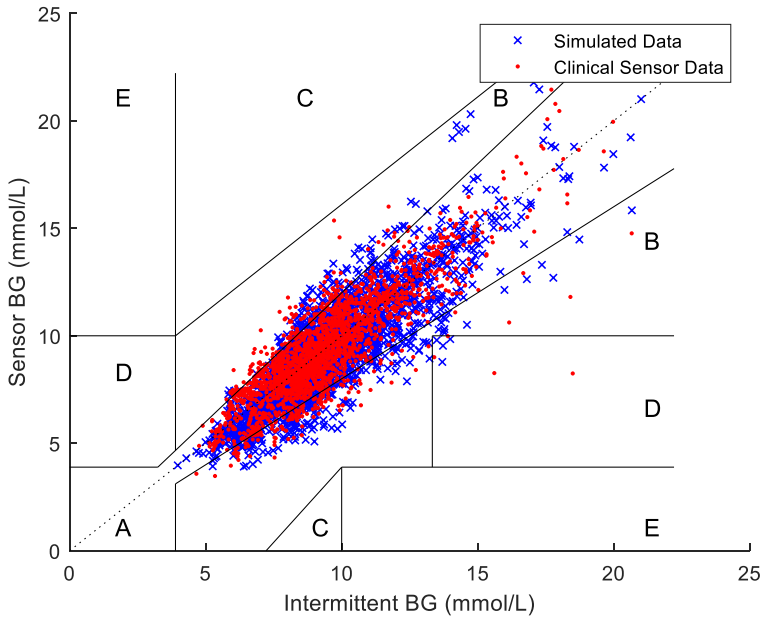


Figure 4.7: Clarke Error Grid plot of the clinical sensor data and the clinical simulated data, both resampled half hourly.

Table 4.4: Percentages of simulated and clinical measurements falling within the zones of the CEG.

Zone	Simulated Data (%)	Clinical Data (%)
A	89.3	82.7
B	10.6	17.1
C	0	0
D	0.1	0.2
E	0	0

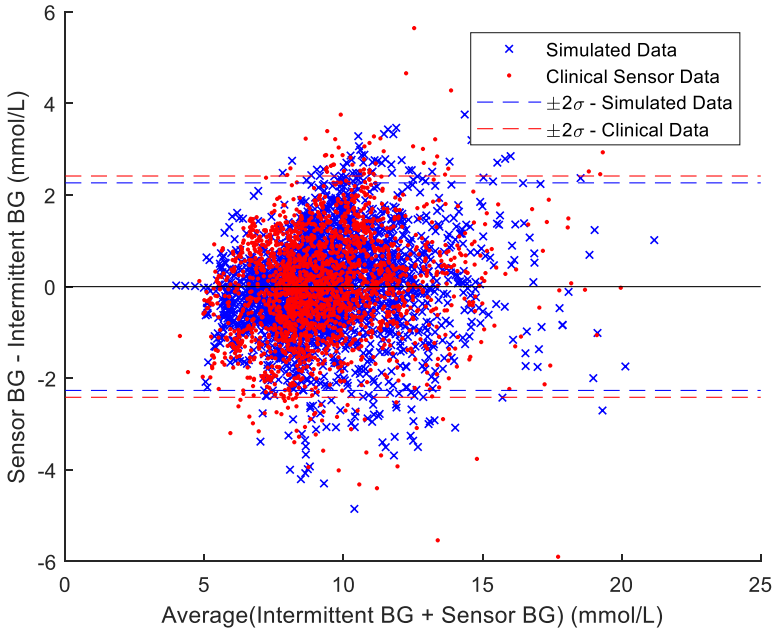


Figure 4.8: Bland-Altman plot of clinical sensor data and simulated clinical data.

Figure 4.9 shows the Trend Compass [167] plot for the clinical sensor data and simulated data. The simulated data matches the clinical data well, indicating the model captures the trend accuracy of the clinical sensor, as well as the point accuracy. The Trend Index, as described in Signal et al. [167], of the clinical sensor and simulated sensor were 10.9° and 11.4° respectively, while the IQR of the theta (θ) values used to evaluate the Trend Index were [3.4° 16.2°] and [4.0° 16.1°] for the clinical data and simulated data, respectively. The similarity between the Trend Indexes and the IQRs of theta values further show that the behaviour of the model is consistent with the clinical data, particularly important for trend simulation which has not been covered in other models. Overall, the model quantitatively represents the sensor behaviour well.

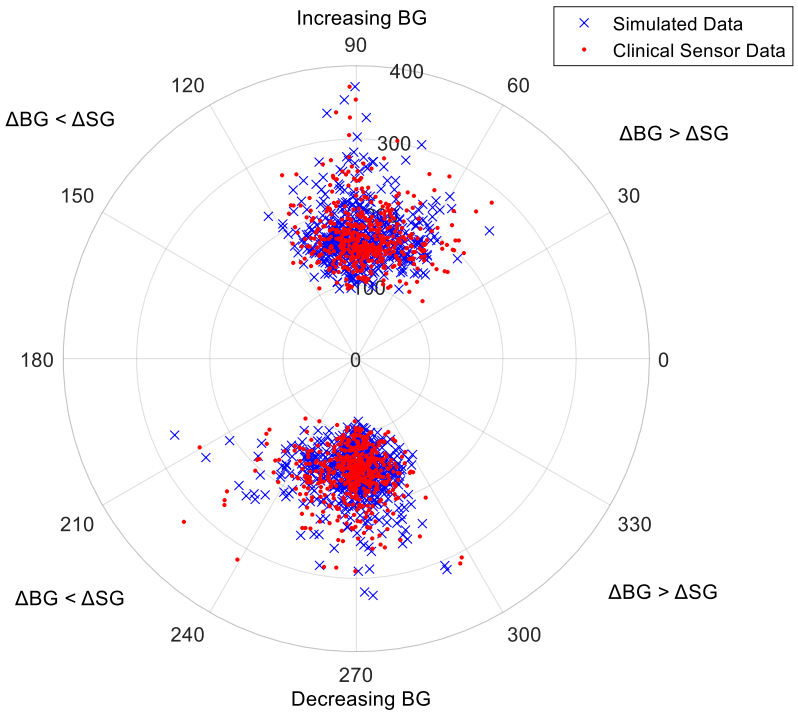


Figure 4.9: Trend Compass plot of clinical sensor data and simulated clinical data.

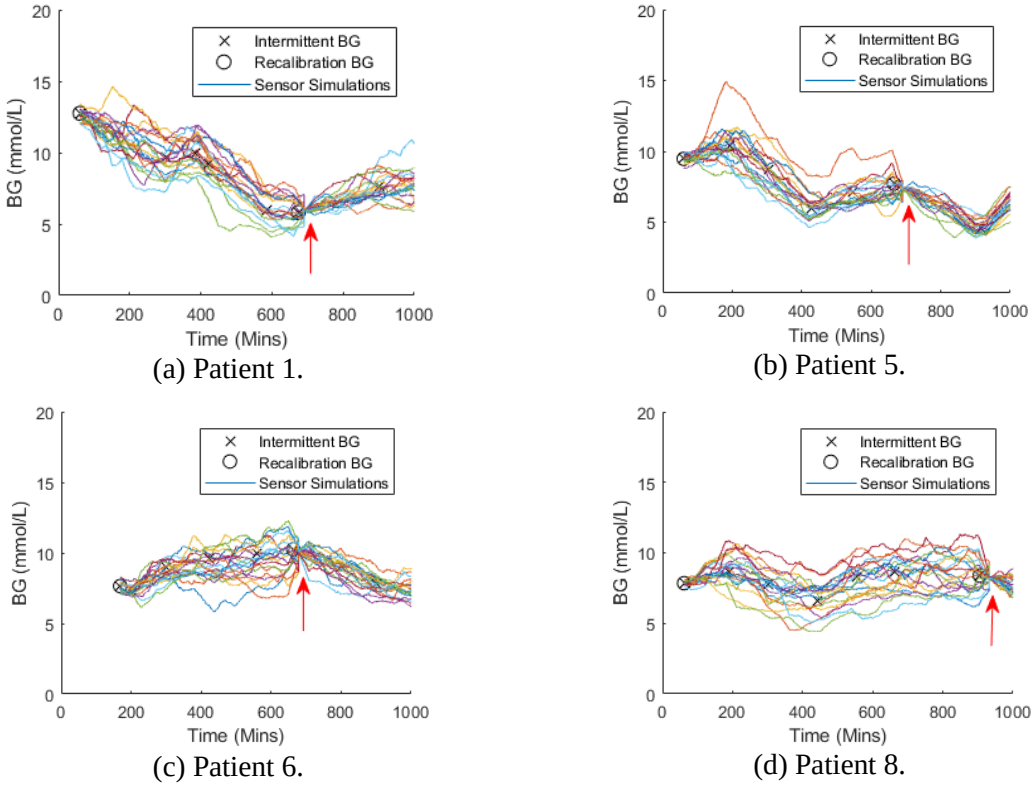
Table 4.5 compares the median clinical SG for each patient, as measured by the CGM sensor at each intermittent BG measurement time, to the average median value from the 100 model simulations of SG. The measured median values for each patient are comparable and differ only slightly, as shown by the percentage error, which had median and IQR range values of 1.2%, -1.1% and 3.1%, respectively. Differences are primarily due to the application of a cohort-model to individual patients, confirming that the model is able to capture the average sensor behaviour over the cohort well. The largest percentage error occurred for Patient 32, which had a clinical sensor reading that was consistently lower than the intermittent BG measurements. This different CGM sensor behaviour could be due to sensor malfunction in the clinical trial, and is not necessarily a failure of the characterisation method or simulation method.

Table 4.5: Median SG for the clinical data and the mean of the median simulated SG value over 100 simulations for each patient.

Patient	Median Clinical SG (mmol/L)	Average Median Simulated SG (mmol/L)	Percentage Error between Clinical Median and Simulated Median (%)
1	9.7	9.3	4.1
2	6.9	6.7	2.9
3	9.4	8.8	6.4
4	10.1	10	1.0
5	8.3	8.2	1.2
6	8.6	8.7	-1.2
7	8.1	8.2	-1.2
8	8.6	8.8	-2.3
9	7.9	7.8	1.3
10	-	-	-
11	8	7.7	3.8
12	9.7	9.6	1.0
13	10.1	10.5	-4.0
14	7.9	8.2	-3.8
15	9.9	10.1	-2.0
16	9.6	9.6	0.0
17	8.1	7.8	3.7
18	9.7	9.5	2.1
19	9.7	9.4	3.1
20	11.1	11.4	-2.7
21	9.4	9.1	3.2
22	-	-	-
23	10.2	10.2	0.0
24	12.9	12.5	3.1
25	8.1	7.8	3.7
26	9	9.1	-1.1
27	9.4	9.3	1.1
28	10.8	10.4	3.7
29	8.4	8.2	2.4
30	8.9	8.6	3.4
31	8.4	8.2	2.4
32	9.5	11.6	-22.1
33	8.6	8.5	1.2
Cohort Median (Range)	9.4 (6.9 – 12.9)	9.1 (6.7 – 12.5)	-

The average global MARD for each simulated patient trace over the 100 sensor simulations, excluding Patients 10 and 22, was 9.6%, where the range of the global MARD of each simulation was from 8.3% to 10.9% over all 31 patients. These values compare well with the clinical global MARD of 9.9%.

Figure 4.10 shows examples of four typical patients individual simulations, with the first 20 of the 100 total simulations plotted for each patient. They each show the multitude of potential sensor traces for a patient given their particular intermittent BG measurements, with a single (arrow) recalibration point included for each patient. Note that each sensor simulation, regardless of the amount of sensor drift at the particular point in time, collapses to the recalibration BG at the recalibration time from the clinical trial, before initiating divergence again through the ξ terms in the model.

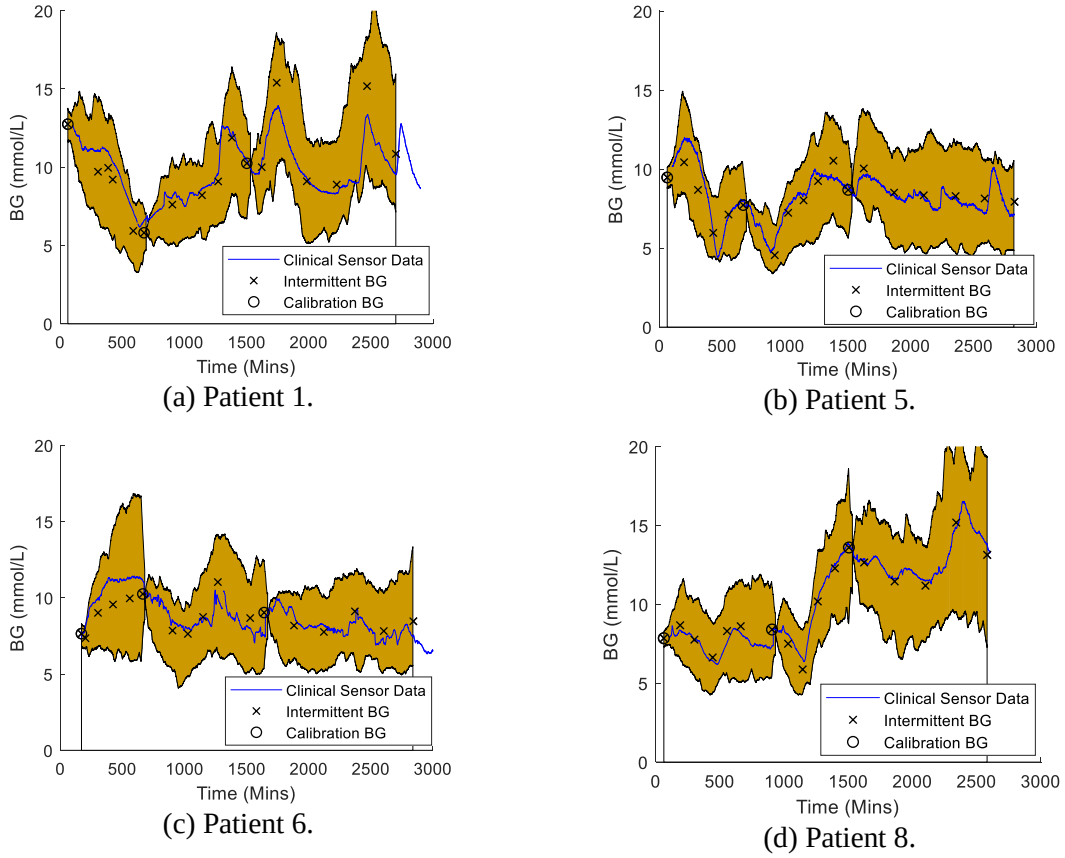


(a) Patient 1. (b) Patient 5.

(c) Patient 6. (d) Patient 8.

Figure 4.10: 20 sensor simulations (out of 100 total) plotted for some example patients, with clinical reference and calibration BG plotted for reference.

Figure 4.11 shows examples of the range of the simulated sensor traces plotted against the clinical CGM sensor data for the same 4 patients in Figure 4.10. These plots show the range of simulated SG compared with clinically measured SG and intermittent BG. This range, or the outlying lines, shows the extremes of the simulated profiles from 100 simulations. The clinical data fell within this simulated range for a large majority of time (> 95% for 88% of patients) or all the time (70% of patients) for most patient profiles.



(a) Patient 1.

(b) Patient 5.

(c) Patient 6.

(d) Patient 8.

Figure 4.11: Clinical CGM sensor data (blue) plotted against the range profile generated with 100 Monte-Carlo simulations (gold).

4.3.3 Convergence Analysis

Figure 4.12 shows the percentage difference distribution plot of the simulated sensor using the re-characterised model of the reduced reference measurement data set. The plot is very similar to the plot of the full sensor model in Figure 4.6, where the minor differences can be explained by the random nature of the CGM sensor trace simulation method. Figure 4.13 also shows the CEG plot of the re-characterised model simulation with reduced reference measurements. The plot is very similar to the plot of the full sensor model in Figure 4.7, while Table 4.6 gives a break down of the percentages of measurement pairs falling within the zones of the CEG plot. Again, the percentages are very similar to the percentages of the full sensor model in Table 4.4, with minor differences also being able to be explained by the random nature of the simulation method. The global MARD for the reduced reference measurement simulation was calculated to be 10.1%, which again is similar to the full sensor model simulation global MARD of 9.6%.

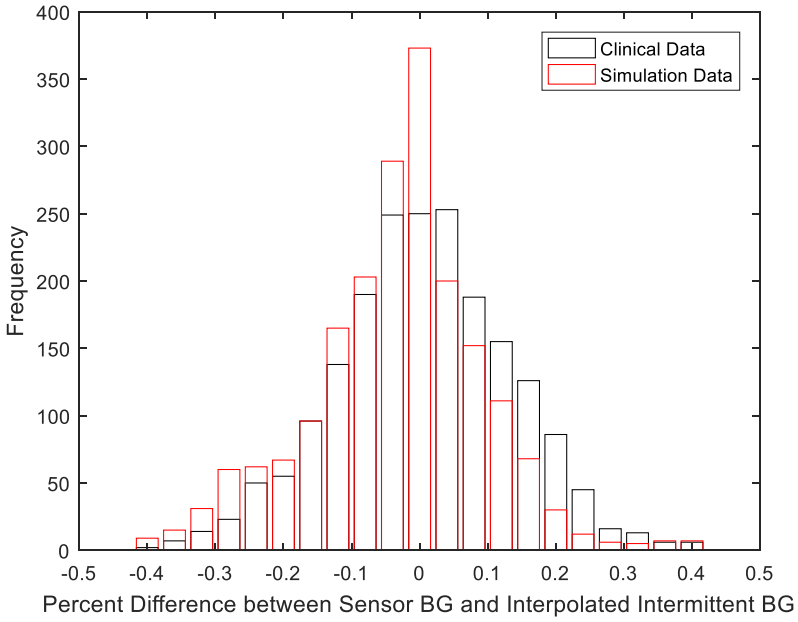


Figure 4.12: Percentage difference distribution plot of the re-characterised model simulation with reduced reference measurements.

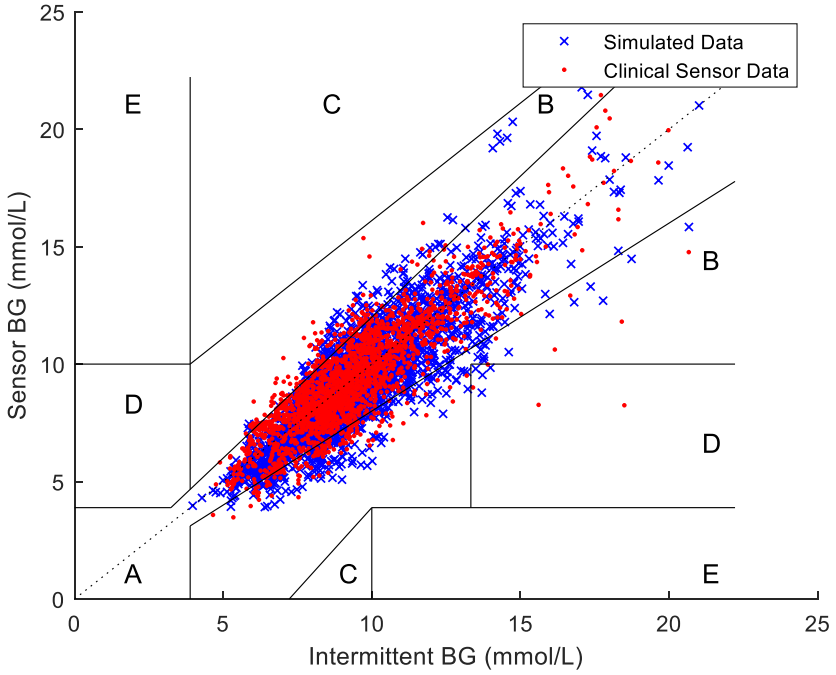


Figure 4.13: CEG plot of the re-characterised model simulation with reduced reference measurements.

Table 4.6: Percentages of simulated and clinical measurements falling within the zones of the CEG.

Zone	Simulated Data (%)	Clinical Data (%)
A	86.1	82.7
B	13.7	17.1
C	0	0
D	0.2	0.2
E	0	0

4.4 Discussion

4.4.1 Sensor Model

Overall, a sensor model was developed from clinical data and simulated measured sensor behaviour well. The MARD, CEG plot and Trend Compass plot for the simulated sensor and the clinical data are very similar (MARD 9.6% vs 9.9%, Trend Index 11.4° vs. 10.9° respectively). Equally, simulated sensor traces were difficult to visually distinguish from clinical data, and clinical data fell within the simulated sensor range for a large majority of time (> 95% for 88% of patients) or all the time (70% of patients) for all patient profiles. Overall, results suggest the method and model are qualitatively and quantitatively able to describe virtually all the observed CGM sensor behaviour, including accurately capturing trend behaviours, which is important for future work testing the sensor in simulation for CGM-based protocol development and optimisation.

The sensor model developed here builds on the sensor model developed by Thomas et al. [182], which uses a much simpler statistical model. The model developed by Thomas et al. was made from Medtronic Sentrino CGM sensor data from an observational pilot study of CGM sensors in patients admitted to the Christchurch Hospital ICU. The key differences between the two models is that the sensor model described in this chapter has been expanded to the lag-2 autoregressive functions to both drift and sensor fluctuations, described as drift and sensor noise respectively in the work presented by Thomas et al., and the additional noise term for each modelled sensor error component.

This method of sensor error characterisation compares well with the method developed by Facchinetti et al. [108], with both methods able to achieve a similarity between the simulated global MARD and the global MARD from clinical studies of the respective devices studied [108, 121, 183]. The method presented here and the method developed by Facchinetti et al. also have the advantage of explicitly and independently characterising sensor drift compared to other prior models [113, 128, 153]. The method presented here could also be perceived as more straight-forward to implement than the method developed by Facchinetti et al., requiring much less data, and also could be used on existing data sets where only data from one CGM device per patient is available, which is currently typical.

Another possible advantage of the method presented here is it may require less reference BG to characterise the error from a sensor and simulate sensor error compared to [108]. The clinical data had reference BG measurements every 1-4 hours, which were then interpolated to carry out the characterisation step. The clinical data Facchinetti et al used required a reference BG measurement every 15 minutes, possibly to enable a more accurate parameter identification of the parameters used in the model. While this high level of clinical data allows precise sensor characterisation, it is highly intensive and costly to gather and would not be available in typical clinical or regulatory trial data. Thus, the model created here can be more easily developed and implemented.

The reproduction of trend accuracy within CGM sensor simulation is an important advantage of this modelling method and has not been tested on other modelling efforts in other publications. Lower trend accuracy in simulation than seen in the clinical sensor can result in more episodes of undetected hypoglycaemia, or a higher rate of false alarms than what would be seen in practice [167]. Conversely, higher trend accuracy in the simulated sensor than in the clinical device would hide the number of missed hypoglycaemic events and give a lower rate of false alarms than what would occur in practice, making the simulation performance seem much better than clinical performance. It is thus a critical feature overlooked by other methods. Both scenarios are mitigated in this case by reproducing the CGM sensor accuracy in the model presented here.

A key result of having a model and generalisable sensor modelling method for a CGM sensor that accurately reproduces point accuracy and trend accuracy, is being able to test the effects of using CGM sensor readings in place of intermittent BG readings during GC in a virtual patient environment. The model can then be used in virtual patient trials to optimise a CGM-enabled GC protocol for any characterised CGM sensor. More importantly, such models, and a generalisable method for making them, also enable the ability to assess what level of CGM performance makes the technology feasible in the ICU to safely improve care and reduce workload.

New developments in CGM devices also have resulted in BG being measured at a higher frequency, such as the device used in this study that is able to measure BG 4 times per minute. A benefit of the model presented here is the ability to capture both the fast and slow random error dynamics of the CGM device. The capturing frequency can be adjusted to match the measurement frequency of other CGM devices that are to have the sensor errors characterised and simulated.

4.4.2 Limitations

One of the limitations of this study is the limited BG < 5.0 mmol/L in the clinical data used to construct the model. As a result, use of this model outside the BG range used to generate it invokes the assumption that sensor behaviour is consistent as a percentage across the BG range, further implying smaller absolute errors at lower BG, and higher absolute errors at higher BG. This assumption of sensor behaviour is consistent with some studies carried out utilising brands of interstitial CGM devices [98, 121, 184]. However, other studies report an increase in MARD at lower BG [122, 185]. Further research would need to target data in this BG range to develop more a reliable sensor model, if necessary, at lower BG.

The linear assumption in the interpolation of BG can also be questioned, as glucose levels could rise or fall faster or slower than the expected linear line. This change depends on changes in patient condition, as well as any changes in insulin and nutrition inputs, which are not known to us in this dataset. In such

conditions, recent work has shown, including times when the input data are better known, that a linear assumption is the best in terms of overall error distributions [186-188].

Another limitation is sensor signal delays. Both the signal filter and the diffusion through the sensor membrane will introduce some delay, although it is likely a smaller signal delay of 10-30 seconds. However, with the relatively high frequency of sensor measurements the sensor makes (4 times per minute) and the relatively long period of times between treatment decisions for which these measurements might be used for in a glycaemic control protocol (0.5-3 hours), this combined signal delay would not be a hugely significant source of error clinically, and thus was not explicitly accounted for in the development of the model. Future work could include an analysis of signal delay to fully evaluate this hypothesis and evaluate its effect on performance in glycaemic control.

A further limitation arises from the clinical sensor data. There were only 1312 hours of recorded data and the median recording period was 39.4 hours. The data that this model was produced from was also the data that this model was tested on, limiting the conclusions of this analyses. More data, particularly from patients that stay for longer than 37.6 hours would allow accurate modelling of long term sensor behaviour, particularly if errors change over time in situ, as occurs with interstitial sensors [120, 122, 145, 189]. Testing the modelled sensor on other sets of data would also help further reinforce this work.

However, it should also be noted that the convergence analysis testing lesser numbers of measurements and model accuracy showed that reducing the reference measurements by 42% before creating the model did not have not significant effect. In particular, the MARD and Clarke Error Grid results were still very close to those of the device in Figures 4.6 and 4.7. Thus, we can also conclude that unless specific dynamics are missing from these patients in this dataset, the number of hours used to create this model is acceptable in its ability to capture the fundamental dynamics

Finally, the analysis was limited by the amount of patient data from the clinical trial. A larger clinical cohort would have allowed more in-depth analysis through cross-validation. In addition, it would enable

more precise characterisation of drift ranges and inclusion of larger drifts in the model would not have skewed results. However, the results from the MARD, CEG plot, Bland-Altman plot and Trend Compass plot indicate that there was enough data to generate a good model that closely captures over 99% of the observed data, since exclusions in sensor modelling eliminated less than 1% of the clinical data.

4.5 Summary

This chapter has shown a CGM sensor can be characterised from patient clinical data using an auto-regressive modelling approach. The method presented here has the benefits of explicitly accounting for sensor drift and requiring far fewer independently sampled blood glucose measures than other methods. Sensor traces can be simulated for BG taken at a clinically realistic rate to create the model. Sensor simulations showed modelled sensor behaviour was very similar to the original clinical data, with very high similarity in MARD, and equally similar Bland-Altman and Clark Error Grid results further validating the model. The novel use of the Trend Compass to validate the trend accuracy reproduction within simulation further showed that the model method is able to accurately capture both point accuracy and trend accuracy. The overall model method is general to any similar sensor and readily extended to interstitial sensors, with or without including interstitial glucose dynamics. It is easily simulated on typical clinical data and thus readily able to be incorporated into proven virtual patients to optimise protocol designs to utilise CGMs in the intensive care unit for glycaemic control.

Chapter 5 CGM STAR

5.1 Background

Continuous glucose monitoring (CGM) technology has improved significantly over the last decade, and has become more prevalent in managing diabetes. New generation CGM sensors [93, 107, 117, 121, 168, 190] with less measurement error than previous iterations [191, 192] have also emerged. These CGM sensors have been used in increasing numbers of studies considering them for glycaemic control (GC) in the intensive care unit (ICU) [80, 95, 129-131, 136, 140, 172, 193-196].

The frequent measures available from CGM sensors offer significant opportunity to monitor and improve the safety and performance of GC without impacting workload. In particular, they can mitigate hypoglycaemia, which can be exacerbated by extended intermittent blood glucose measurement intervals [49, 102, 173], and may be an important consideration moving forward in the assessment of the merits of glycaemic control in the ICU.

In particular, the hypoglycaemia and variable control that plague many studies often occurs where infrequent blood glucose (BG) measurements combine with highly variable changes in patient response to care and complex protocols [103, 128, 148, 197]. The increased temporal resolution CGM devices provide can monitor real-time BG trends. This monitoring, in turn, allows more rapid treatment response to highly dynamic changes in patient condition and glycaemia [24, 32], by modifying insulin and/or nutrition delivery to avoid hypoglycaemic excursions, where this variability is well-quantified in many cohorts [51-53, 156].

Both intravascular (IV) and subcutaneous CGM sensors have been used for measurement to guide GC, either through use of the CGM sensor readings in a GC protocol [129-131, 136, 140, 172, 193, 194], or to observe trends and detect hyper/hypoglycaemia while retaining a GC protocol with intermittent

measures [80, 95, 195, 196]. Results have been mixed, with some studies showing good GC results with high percentages of measurements in target bands and low numbers of hypoglycaemic incidents [80, 95, 130, 196]. However, the studies with these positive outcomes tend to be small pilot trials on relatively more stable patients [80, 95, 196]. The mixed results can also be attributed to GC performance being a product of CGM sensor quality, determined by random error and sensor drift, the interaction of the CGM sensor with the specific GC protocol, and the quality of the GC protocol itself.

CGM devices can reduce nursing GC related workload via automation reducing blood sampling requirements [128-130]. However, the increased temporal measurement resolution CGM devices provide is still somewhat outweighed by the larger point accuracy errors in these devices due to sensor drift, bias, and noise [107, 108, 113, 128, 153, 164-166]. Hence, there is significant need to better integrate CGM sensor dynamics into GC protocols to maximise their utility. This opportunity is most apparent with model based GC and virtual patient driven care [34, 51, 54, 62, 151, 172, 179, 198].

Interstitial CGM devices have been modelled in the past for testing in virtual environments [107-109, 113, 128, 153, 182, 199]. However, the previous models have only been used to model CGM measurements during closed-loop control of diabetes in the much less dynamic non-critical care environment [175]. Only one prior study has completely modelled detailed CGM sensor traces in a critical care environment [107].

This chapter presents the results of using the IV CGM sensor model presented in Chapter 4 to design and optimise GC with the established, proven STAR GC protocol [34, 54, 61, 62, 200, 201]. The STAR protocol framework is adapted to utilise the increased temporal resolution of a CGM device within a typical ICU intervention framework. Clinically validated virtual trials are used to test and optimise this combination *in-silico*, and delineate the trade-off of sensor error, safety, performance and workload.

5.2 Methods

Clinically validated virtual trial simulations [149-151] are used to assess the performance, safety and nurse workload associated with using CGM sensors to guide GC in a critical care environment. The process is summarised in Figure 5.1, where virtual patients are generated from clinical data using the ICING (Intensive Care Insulin-Nutrition-Glucose) model [146] and virtual trial simulation is used to simulate glycaemic outcomes. In this chapter, virtual trials were performed with a CGM-enabled version of the STAR insulin-nutrition protocol and the IV CGM sensor model developed in Chapter 4. Multiple cases of CGM-enabled STAR virtual trials were run to determine the best combination of safety features using Monte-Carlo simulation.

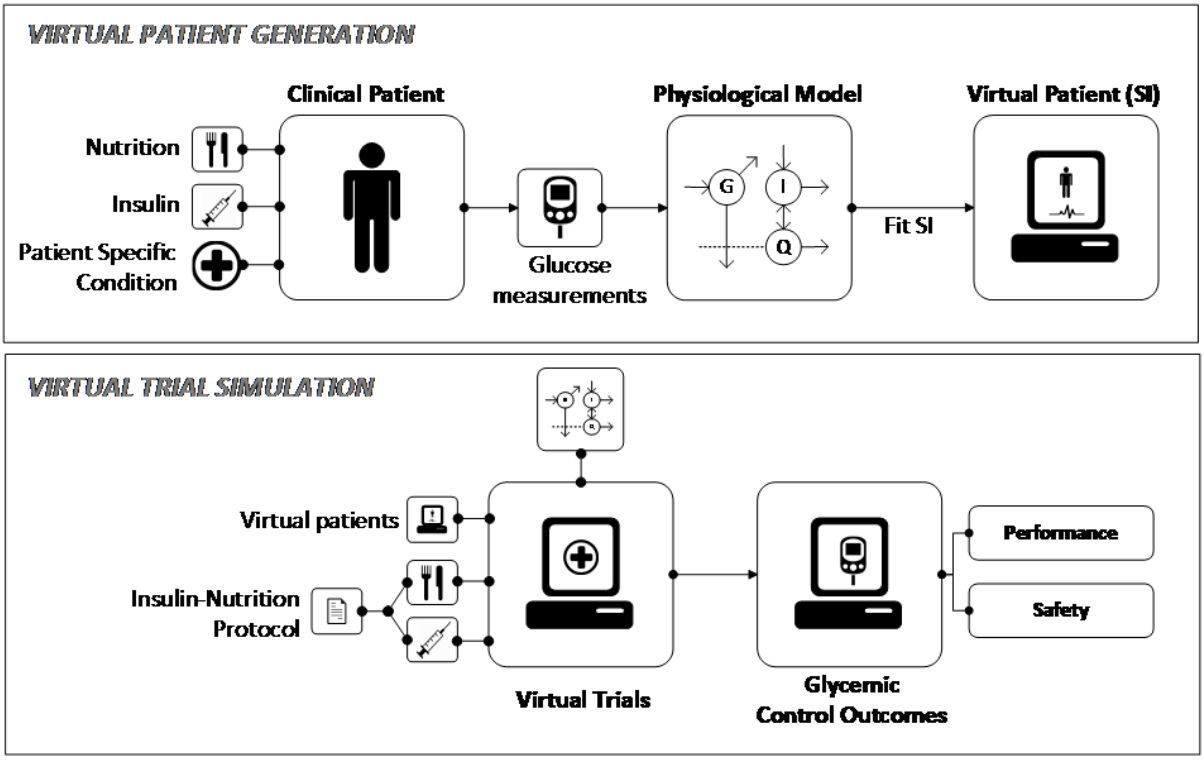


Figure 5.1: Summary of the virtual patient generation and virtual trial simulation process.

5.2.1 Virtual Patient Generation

Virtual patients consist of a time-varying insulin sensitivity (SI) profile identified from clinical BG, nutrition, and insulin data [202-204]. This process, summarised in the top box of Figure 5.1, has been clinically validated on independent data, in clinical use, and in prediction of trial outcomes across

multiple cohorts [62, 149, 150, 152]. A virtual trial uses this SI profile to simulate BG outcomes for different treatment inputs [54, 62, 205].

5.2.1.1 Clinical Data

Virtual patients were generated from a clinical data cohort of 236 patients who received insulin therapy for more than 24 hours under the STAR protocol [34] from June 2011 to May 2015. Patient demographic data is shown in Table 5.1. Additional data and information for these patients can be found in Stewart et al. [34].

Table 5.1: Patient demographic data for the 236 clinical patients used to generate virtual patients.

STAR Cohort		
Total Patients	236	
Age (Years)	64.5 [53 - 72]	
% Male	67.4%	
Length of Stay (Days)	9.8 [3.9 - 19.2]	
Length of Glycaemic Control (Days)	3.7 [2.0 - 7.1]	
APACHE II Score	21 [17 - 26]	
APACHE II Risk of Death	34.1% [18.4% - 54.9%]	
APACHE III Diagnosis:		
Operative	Num. Patients	%
Cardiovascular	35	15
Respiratory	5	2
Gastrointestinal	10	4
Neurological	7	3
Trauma	10	4
Other (Musculoskeletal, renal/genitourinary)	5	2
Non-operative	Num. Patients	%
Cardiovascular	33	14
Respiratory	60	25
Gastrointestinal	15	6
Neurological	18	8
Trauma	17	7
Other (Sepsis, metabolic, renal)	21	9

Data are median [IQR] where applicable.

5.2.1.2 Model

Model based SI is identified using the clinically-validated ICING metabolic model [146]:

$$\dot{G}(t) = -p_G G(t) - S_I(t)G(t)\frac{Q(t)}{1+\alpha_G Q(t)} + \frac{P(t) + EGP - CNS}{V_G} \quad (5.1)$$

$$\dot{Q}(t) = n_I(I(t) - Q(t)) - n_C\frac{Q(t)}{1+\alpha_G Q(t)} \quad (5.2)$$

$$\dot{I}(t) = -n_K I(t) - n_L\frac{I(t)}{1+\alpha_I I(t)} - n_I(I(t) - Q(t)) + \frac{u_{ex}(t)}{V_I} + (1 - x_L)\frac{u_{en}(G)}{V_I} \quad (5.3)$$

Where Table 5.2 contains all the variable and parameter definitions and values for the ICING model.

Table 5.2: ICING model variables and parameter definitions and values.

ICING model variables	
Blood glucose (G)	
Plasma insulin (I)	
Interstitial insulin (Q)	
Insulin mediated cellular glucose uptake (SI)	
Exogenous glucose appearance from parenteral and enteral sources (P)	
Exogenous insulin (u_{ex})	
Endogenous insulin (u_{en})	
ICING model constant parameters	**Model values**
Endogenous glucose production (EGP)	1.16 mmol/min or 20.88 mg/dl
Central nervous system uptake (CNS)	0.3 mmol/min or 5.4 mg/dl
Non-insulin mediated uptake (p_G)	0.006 min^{-1}
Kidney insulin clearance (n_K)	0.0542 min^{-1}
Liver insulin clearance (n_L, x_L)	0.1578 min^{-1}, 0.67
Insulin diffusion to the interstitial space (n_I)	0.0060 min^{-1}
Interstitial insulin degradation (n_C)	0.0060 min^{-1}
Glucose distribution volume (V_G)	13.3 L
Insulin distribution volume (V_I)	3.15 L

SI is the only parameter identified on a patient specific basis. It varies hourly with time, and is identified from clinical data using integral based fitting [202-204]. It is this time varying SI profile that defines a virtual patient. Thus, virtual patients are based off clinical data and clinically observed changes in insulin sensitivity, and are metabolic "clones" of the actual patient [151].

5.2.2 Virtual Trial Simulation

Virtual trials are carried out using virtual patients, a CGM sensor model, and the STAR protocol. Multiple simulation cases are run to optimise the CGM-enabled protocol and performance metrics for safety, performance and workload allowed for the evaluation of each simulation case.

5.2.2.1 CGM Sensor Model

The IV CGM sensor model is presented in detail in Chapter 4. However, in summary, two main sensor errors simulated are sensor drift (*Drift*) and sensor fluctuations (*SensorFlux*) via auto-regressive (AR) processes. Intermittent clinical BG_{IM} measurements are used as a base from which to simulate CGM sensor measurements for use in virtual trials [149, 179-181]. Clinical BG measurements are interpolated at half hourly intervals from the last measurement such that:

$$BG_{IM/30\ min} = interp(BG_{IM})|_{t=0:30:t_{end}} \tag{5.4}$$

Drift is applied to each half hourly interpolated BG such that:

$$BG_{base(t=0:30:t_{end})} = BG_{IM/30\ minutes} * (1 + Drift) \tag{5.5}$$

Drift is as described in Chapter 4, where it is defined:

$$Drift_{n+1} = \alpha_d + \beta_d * Drift_{n-1} + \gamma_d * Drift_n + \xi_d \tag{5.6}$$

Where α_d, β_d and γ_d are AR parameters, and ξ_d is a random variable. This BG_{base} is further interpolated to provide 4 measurements per minute, the same rate of measurement as the clinical CGM sensor, such that:

$$BG_{base/0.25\ minutes} = interp(BG_{base})|_{t=0:0.25:t_{end}} \tag{5.7}$$

The simulated sensor output is then defined:

$$BG_{new\_sensor(t=0:0.25:t_{end})} = BG_{base/0.25\ minutes} * (1 + SensorFlux) \tag{5.8}$$

Where *SensorFlux* has been defined as in Chapter 4 such that:

$$SensorFlux_{n+1} = \alpha_{sf} + \beta_{sf} * SensorFlux_{n-1} + \gamma_{sf} * SensorFlux_n + \xi_{sf} \quad (5.9)$$

Where α_{sf}, β_{sf} and γ_{sf} are also AR parameters, and ξ_{sf} is a random variable.

The model parameter values of α, β and γ for both drift and sensor fluctuations are presented in Table 5.3. The sensor model has been validated against clinical trial data in Chapter 4 and [107]. Equations (5.4)-(5.9) thus define the means of generating realistic CGM measures from intermittent clinical BG data in virtual trials.

Table 5.3: Auto-regressive (AR) model parameters for drift and sensor fluctuations.

AR pass	Key characteristic	α	β	γ	ξ median	ξ max (absolute)
1	Drift	-0.00147	-0.07152	0.9698	-0.0023	0.2486
2	Sensor Fluctuations	-6.0e-6	-0.261	1.261	-2.66e-06	0.01

5.2.2.2 STAR Protocol

STAR (Stochastic TARgeted), is a computerised model-based GC protocol utilising stochastic models to forecast metabolic variations and glycaemic outcomes for a given insulin and nutrition intervention [34, 54, 61, 62]. Future SI values are forecast from current identified SI values using stochastic models built from cohort data [52, 146]. These forecast SI values can be used to generate probable BG outcome distributions for any given insulin and nutrition intervention. STAR overlaps this distribution of potential BG outcomes with the clinically targeted 4.4 – 8.0 mmol/L range, using the 5th percentile of likely BG outcomes to ensure a 5% maximum risk of BG < 4.4 mmol/L. Insulin and nutrition treatments are selected to meet this goal, with a secondary goal of maximising nutrition. The stochastic model and corresponding SI and BG forecasts used to determine these insulin and nutrition treatments are shown in Figure 5.2.

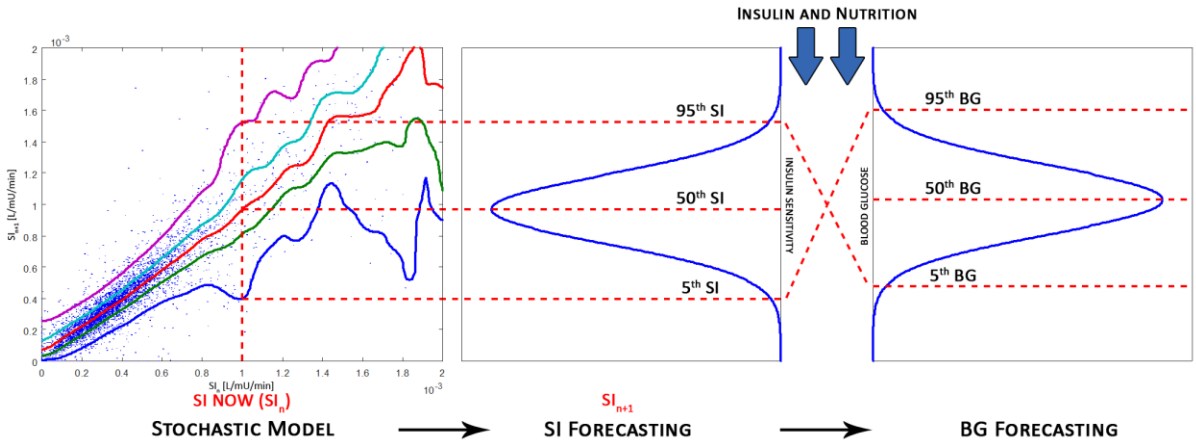


Figure 5.2: Stochastic model shows the current SI plotted against the percentage likelihood of future SI. The stochastic model is used to determine the SI forecast and corresponding BG forecast.

In clinical use, BG is measured 1-hourly outside the 4.4 – 8.0 mmol/L range, and STAR offers 2 and 3 hour options for patients in the target range. In simulation, the maximum available measurement interval is always chosen. However, clinically it is left to staff discretion, resulting in slightly higher, voluntarily chosen measurement rates in some units [34].

5.2.2.3 CGM-enabled STAR

The CGM sensor model of Equations (5.4)-(5.9) is used to replace intermittent 1-3 hourly BG measurements in STAR. To simulate automation with the CGM sensor, interventions are calculated hourly using the modelled CGM sensor glucose (SG) value, unless a recalibration measure is taken, in which case this recalibration value is used instead. A typical clinical workload revolves around interventions at 1-3 hourly intervals, and so hourly CGM measures reduces required blood draws and thus clinical workload, while potentially improving control with more frequent adjustment. Hourly GC control can be very tight and automated CGM measures enable this approach for clinical practice, where it would otherwise be infeasible due to workload [71, 104].

CGM recalibrations reset SG to the calibration BG value assuming a point-to-point match [107]. Recalibrations are triggered when any of: (1) a specified glucose guard rail is crossed; (2) there is a rapid change in CGM measured blood glucose; or (3) 8 hours has passed since the last recalibration [93]. There is thus a minimum possible workload of 3 measures per day, unless conditions (1)-(2) are triggered.

Guard-rails based on CGM measures are used to trigger calibration measurements ensuring higher sensor accuracy at areas of higher risk. Pairs of guardrail values were used, comprised of upper and lower BG thresholds for recalibration. Guard-rails of (4.0 mmol/L, 8.0 mmol/L) and (4.4 mmol/L, 8.0 mmol/L) were tried, defining the intermediate BG range between hypo- and hyperglycaemia [25, 26, 65]. Any single SG measurement outside this range triggered a recalibration measure, as well as a STAR insulin-nutrition intervention change based on the recalibration blood glucose measurement.

A rolling window monitoring changes in SG is also used to trigger recalibration measures. The rolling window looks at the change in SG measurement over a specified time frame, and this window rolls forward as more SG measures are available. If the magnitude of change in SG over this window exceeded a specified threshold, a recalibration is triggered and a treatment recommended using the recalibration measure. Sensitivity of GC performance to the interval and magnitude of SG change in these rolling windows was evaluated. Rolling window intervals of 20, 30, and 40 minutes were tested, as well as absolute changes in BG of 0.5, 1.0 and 1.5 mmol/L over these intervals. The minimum time allowed between consecutive recalibrations was set to 1 hour to avoid excessive increases in workload that would not significantly affect sensor performance.

In the Christchurch Hospital ICU, insulin therapy is delivered as boluses every 1-3 hours. In this study, insulin was administered via insulin infusions, a delivery method more commonly used elsewhere, and easier to administer in an automated ‘closed loop’ manner. In an automated closed loop, full automation

would not require a nurse to be present to program boluses or infusions, thus further lowering clinical workload. This study examines the impact of CGM sensors on safety and performance, as a step towards further or complete automation.

5.2.2.4 Virtual Trials Analysis

A simulation of clinical STAR, in which BG is measured intermittently every 1-3 hours and insulin is delivered in the form of boluses (Case B), was run to compare to the clinical STAR data (Case A), to show the difference between clinical and virtual patients in addition to the impact of always choosing the longest measurement intervals. STAR with 1 hour BG measurements and insulin infusions (Case C) was simulated to provide a baseline for comparison, where the impractical measurement rate of 24 BG measurements per day represents an upper bound of what performance might be expected with a CGM device measuring hourly. Table 5.4 lists all virtual trials utilising the IV CGM model. Three Monte-Carlo simulations are carried out per virtual trial in Table 4, bringing the total number of virtual patient sensor trace outputs to 708 for each virtual trial. In particular, because the CGM model has a random component, no two or more runs of a given virtual patient have the same CGM trace. Multiple runs and many patients alleviates any issue of outlier behaviour affecting results.

Cases 1 and 2 were run to optimise the BG levels of the guard rails for GC performance, safety and workload. Case 3 repeats Case 1, but allows for intervention interval length to extend up to 3 hours to assess this parameter in a semi-automated case where changing interventions hourly has a workload cost. A total of 9 simulations (Cases 4-12) were carried out to evaluate the effect of changing the parameters of the rolling window. Finally, recalibration periods were varied from 6-12 hours (Cases 13-14) to assess robustness to this metric.

Table 5.4: List of simulation cases run using CGM sensor model.

Case Number	Guard rails (mmol/L)	Recalibration Period (h)	Rolling Window	
			Window Length (t) (Minutes)	ΔBG with BG between guardrails (4.4 and 8.0 mmol/L)
1	4.0 and 8.0	8	-	-
2	4.4 and 8.0	8	-	-
3 (1-3 hour interventions)	4.0 and 8.0	8	-	-
4	4.4 and 8.0	8	40	±1.0 mmol/L
5	4.4 and 8.0	8	30	±1.0 mmol/L
6	4.4 and 8.0	8	20	±1.0 mmol/L
7	4.4 and 8.0	8	40	±0.5 mmol/L
8	4.4 and 8.0	8	30	±0.5 mmol/L
9	4.4 and 8.0	8	20	±0.5 mmol/L
10	4.4 and 8.0	8	40	±1.5 mmol/L
11	4.4 and 8.0	8	30	±1.5 mmol/L
12	4.4 and 8.0	8	20	±1.5 mmol/L
13	4.0 and 8.0	6	-	-
14	4.0 and 8.0	12	-	-

5.2.2.5 Performance Metrics

Hourly resampled measurements are used so direct comparisons can be made between cohorts with different measurement and intervention intervals. For each virtual trial, the following performance indicators are used:

- **Safety** was measured as the percentage of patients who had a severe hypoglycaemic episode (BG < 2.22 mmol/L), as well as the percentage of resampled BG < 4.0 mmol/L.
- **Performance** was measured as the percentage of hourly resampled measurements in the 4.4 – 7.0 mmol/L and the 4.4 – 8.0 mmol/L target bands [25, 26, 65].
- **Workload** was measured by the number of manual blood glucose measures per day.

Each virtual trial in which the CGM sensor is employed also recorded the SG results, with added sensor errors, alongside the true (simulated) BG results to evaluate the impact on performance in true BG against what would be seen as 'true' in a clinical setting where only SG would be known. Thus, all of the performance metrics are measured using true BG to reflect the actual results of each trial, while SG

results are provided to evaluate how the results would have been reported clinically based off observed CGM data including errors. This comparison shows the impact of any sensor errors on perceived versus true performance, which is only possible in virtual trials and may be significant in sensors with larger errors and/or drift.

5.3 Results

5.3.1 Virtual Trials Results

Table 5.5 presents baseline clinical (Case A) and virtual trial results for STAR without CGM devices (Case B). The clinical and simulation results are broadly similar, with discrepancies due to the difference in measurement interval, where clinical staff are free to select 1-3 hours between measurements [61, 62], but in simulation the maximum available is selected. As a result, the main difference between the results was the simulated clinical STAR had a higher percentage of BG measurements in the 4.4 – 7.0 mmol/L band (76.9% vs 62.8%), but a more similar number in the 4.4 – 8.0 mmol/L range (89.0% vs 83.7%). Similarly, the incidence of mild hypoglycaemia was slightly increased in simulation (0.5% vs 1.1% BG < 4.0 mmol/L), while the percentage of patients who experienced a severe hypoglycaemic event (BG < 2.22 mmol/L) were similar and low (1.27 vs 0.85% for simulated and clinical respectively). Thus, Case A and Case B are not drastically different, especially in per-patient results, to say the virtual cohort is a good representation of the clinical cohort.

The last column in Table 5.5 shows Case C, the best case STAR controller results using 1 hour interventions and insulin infusions. GC performance was slightly improved compared to simulated clinical STAR in Case B (80.0% vs 76.9% in the 4.4 – 7.0 mmol/L range, 89.1% vs 89.0% in the 4.4 – 8.0 mmol/L range), while workload was significantly increased (24 interventions per day vs 12.2 in the clinical STAR simulation), as expected. This result is the baseline best case for comparing CGM-enabled STAR results, and also indicates the decreasing gain in performance found with increasing measurement rates.

Table 5.5: Clinical STAR results, simulated clinical results and simulated 1 hour STAR infusions.

	Clinical 1-3 Hr STAR Bolus (Case A)	Simulated Clinical 1-3 Hr STAR Bolus (Case B)	Simulation 1 Hr STAR Infusions (Case C)
Whole Cohort Statistics			
Number of patients	236	236	236
No. BG Measures	11001	10622	20972
No. interventions per day	12.7	12.2	24
Percentage of patients who had at least one measurement < 2.22 mmol/L (No. of patients)	0.85 (2)	1.27 (3)	0.85 (2)
Hourly Resampled Statistics			
% BG within 4.4 - 7.0 mmol/L	62.8	76.9	80.0
% BG within 4.4 - 8.0 mmol/L	83.7	89.0	89.1
% BG within 8.0 - 10 mmol/L	11.4	6.6	6.2
% BG < 5.0 mmol/L	4.4	8.6	13.1
% BG < 4.0 mmol/L	0.51	1.08	1.05
Median BG [IQR] (mmol/L):	6.6 [6.0 – 7.4]	6.1 [5.6 – 6.8]	5.8 [5.3 – 6.6]
Median insulin dose [IQR] (units/hour):	2.5 [2.0 - 3.5]	3.0 [2.0 - 4.5]	4.0 [2.0 - 6.0]
Per-patient median dextrose rate for those fed [IQR] (g/hour):	5.5 [4.4 - 6.7]	5.8 [4.5 - 6.7]	5.9 [4.7 - 6.7]

Table 5.6 presents virtual trial results for Cases 1-4 in Table 4. Each virtual trial presents results in terms of 'true' underlying BG and observed SG in two columns. Performance and safety for all trials in Table 5.6 are broadly similar to the baseline results for Case C in the last column of Table 5.5, but with significantly reduced workload.

In Table 5.6, the number of recalibration, and thus manual BG, measurements per day ranged from 3.0-3.6 per day. The percentage of true BG measurements within the 4.4 – 7.0 mmol/L band ranged from 71.6% to 77.1%, while the percentage of true BG measurements within the 4.4 – 8.0 mmol/L band ranged from 87.0% to 88.1%. The percentage of true BG < 4.0 mmol/L ranged from 1.30% to 1.37%.

Additionally, the percentage of patients who experienced a hypoglycaemic episode ranged from 0.71% (5 of 708 patients) to 1.41% (10 of 708 patients). The last column of Table 5.6 presents the results typical of the virtual trials in which a guard rail and a rolling window were employed.

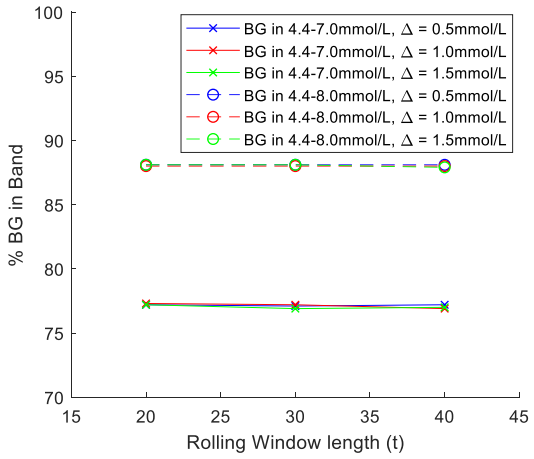
All of the different rolling window definitions (Cases 4 - 12) had very similar virtual trial results, so only one set of full results is shown for brevity in Table 5.6. Figure 5.3 shows the similarity in these 9 results, with the number of recalibrations per day ranging from 3.5-4.1, percentage BG within 4.4 – 7.0 mmol/L band ranging from 76.9 – 77.3%, percentage BG within 4.4 – 8.0 mmol/L band ranging from 87.9 – 88.1%, and percentage BG < 4.0 mmol/L band ranging from 1.25 – 1.33%.

Figure 5.4 shows the virtual trials testing recalibration period (Cases 13-14) for robustness and optimisation. Figure 5.4(a) shows the percentage of BG measurements within the thresholds, and shows no change in percentage BG in band as reported by the sensor and no change in median BG (Table 5.7). Figure 5.4(b) compares the percentage of BG measurements less than the hypoglycaemic thresholds of 2.22 mmol/L and 4.0 mmol/L between the recalibration periods, while Figure 5.4(c) compares the nurse workload and patient safety. Small degradations in performance and safety are observed versus larger reductions in workload, defining a clear trade-off for clinicians.

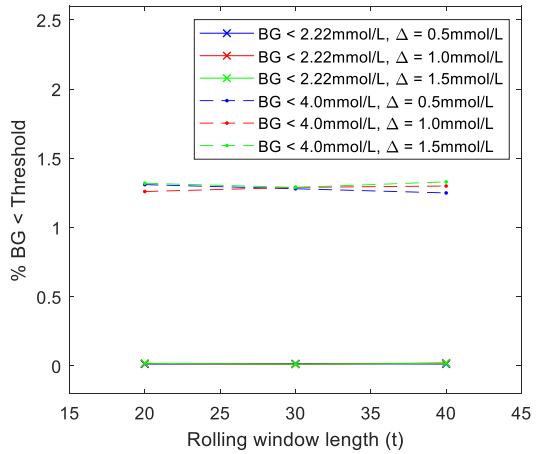
Table 5.6: Results from virtual trials utilising guard-rails, with one representative utilisation of the rolling window.

	Case 1: Guard Rail 4.0-8.0 mmol/L, minimum 8 Hr Recalibration, Hourly Interventions		**Case 2: Guard Rail 4.4-8.0 mmol/L, minimum 8 Hr Recalibration, Hourly Interventions**		**Case 3: Guard Rail 4.0-8.0 mmol/L, Minimum 8 Hr Recalibration, 1-3Hr Interventions**		**Case 4: 40 min rolling window, ΔBG of -1.0 mmol/L with BG < 4.4 mmol/L, ΔBG of +1.0 mmol/L with BG > 8.0 mmol/L, Hourly Interventions**	
Whole Cohort Statistics								
	True BG	SG	True BG	SG	True BG	SG	True BG	SG
Number of patients	708		708		708		708	
Total hours	62520		62566		62892		62553	
No. recalibrations	8603		9186		7944		9359	
No. recalibrations per day	3.3		3.5		3.0		3.6	
Percentage of patients who had at least one measurement < 2.22 mmol/L (No. of patients)	1.13 (8)	1.27 (9)	1.41 (10)	1.41 (10)	0.71 (5)	1.13 (8)	1.41 (10)	2.26 (16)
Hourly Resampled Statistics								
% BG within 4.4 - 7.0 mmol/L	76.6	77.4	77.1	78.2	71.6	73.2	76.9	78.6
% BG within 4.4 - 8.0 mmol/L	87.9	88.4	88.1	89.0	87.0	87.2	88.0	89.1
% BG within 8.0 - 10 mmol/L	6.5	5.5	6.4	5.6	7.7	6.9	6.4	5.5
% BG < 5.0 mmol/L	13.3	16.0	13.5	14.8	10.1	11.6	13.8	15.1
% BG < 4.0 mmol/L	1.37	1.31	1.30	1.49	1.37	1.48	1.30	1.56
% BG < 2.22 mmol/L	0.019	0.024	0.017	0.022	0.014	0.016	0.021	0.027
Median BG [IQR] (mmol/L):	6.0 [5.4 - 6.8]	5.9 [5.3 - 6.7]	6.0 [5.4 - 6.7]	5.9 [5.3 - 6.69]	6.3 [5.6 - 7.0]	6.2 [5.5 - 6.9]	6.0 [5.4 - 6.7]	5.9 [5.3 - 6.7]
Median insulin dose [IQR] (units/hour):	4.0 [1.5 - 8.0]		4.0 [1.5 - 8.0]		3.5 [1.7 - 5.6]		4.0 [1.5 - 8.0]	
Per-patient median dextrose rate for those fed [IQR] (g/hour):	5.5 [4.6 - 6.6]		5.5 [4.5 - 6.6]		5.2 [4.1 - 6.1]		5.5 [4.7 - 6.6]	

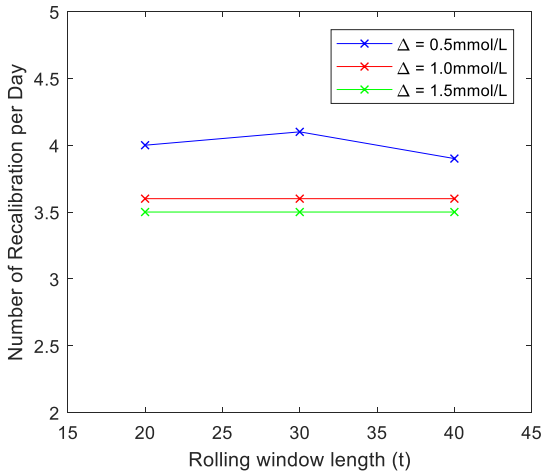
Rolling window parameter testing



(a) Comparing the percentage of BG in band for each of the rolling window lengths.



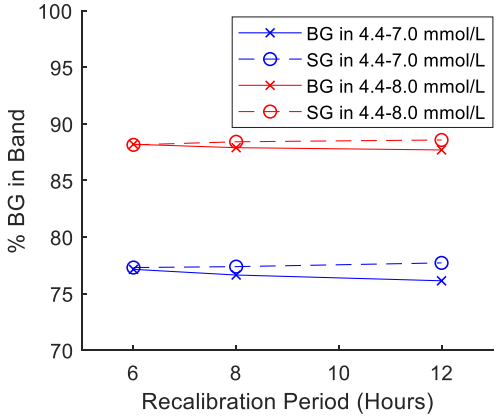
(b) Comparing the percentage of BG less than the threshold for each of the rolling window lengths.



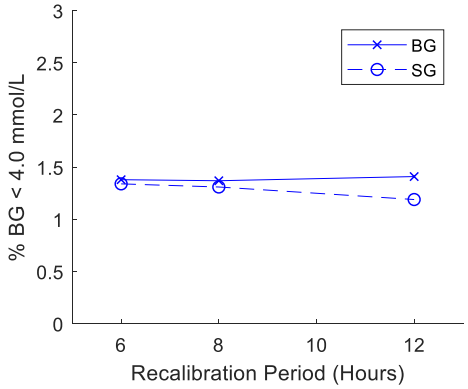
(c) Comparing rolling window length with nurse workload.

Figure 5.3: Graphical results from varying the parameters of the rolling window, with a guard rail of 4.4-8.0 mmol/L.

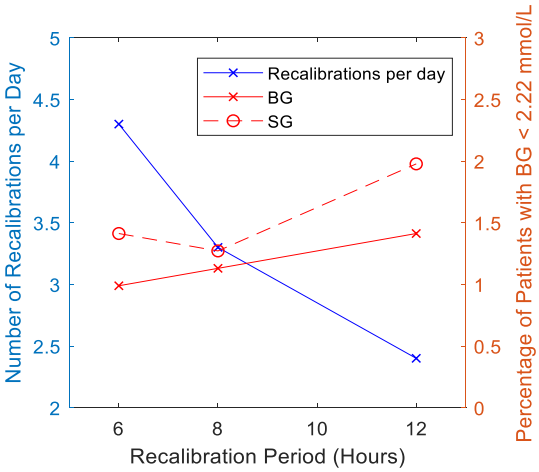
Recalibration robustness trials



(a) Comparing the percentage of BG in band for each of the recalibration periods. 'Real' BG is shown with solid lines, SG is shown with dotted lines.



(b) Comparing the percentage of BG less than the thresholds for each of the recalibration periods. Again, 'real' BG is shown with solid lines, SG is shown with dotted lines.



(c) Comparing recalibration period with nurse workload and patient safety.

Figure 5.4: Graphical results from the testing of varying recalibration period with a guard rail of 4.0-8.0 mmol/L.

Table 5.7: BG, insulin and dextrose results from varying recalibration period.

Recalibration Period (Hours) (Case No.)	Median BG (mmol/L) [IQR] (resampled)		Median Insulin (U/hour) [IQR] (resampled)	Median Dextrose (g/hour) [IQR] (resampled)
	BG	SG		
6 (Case 13)	6.0 [5.4 - 6.7]	5.9 [5.3 - 6.7]	4.0 [1.5 - 8.0]	5.5 [4.6 - 6.6]
8 (Case 1)	6.0 [5.4 - 6.8]	5.9 [5.3 - 6.7]	4.0 [1.5 - 8.0]	5.5 [4.6 - 6.6]
12 (Case 14)	6.0 [5.4 - 6.8]	5.9 [5.3 - 6.9]	4.0 [1.5 - 8.0]	5.6 [4.6 - 6.6]

5.4 Discussion

5.4.1 Sensor Use in Simulated GC

In general, replacing the more accurate intermittent BG measurements of STAR with CGM sensor measurements significantly reduced workload with clinically small to negligible effects on performance and safety. This trade-off is what clinicians must consider before fully implementing CGM sensors in GC. It is based on the underlying sensor drift, bias and error of this CGM [107], where lower errors and drift might improve these results if found via a different or improved sensor.

For Case 1, the number of measurements required per day was 3.3, including recalibrations, an improvement in workload of 74% compared to the 12.7 for the clinical STAR Case A, and 73% for its virtual trial Case B. Utilising CGM sensor measurements resulted in similar performance to simulated STAR, Case B, (76.6% vs 76.9% in the 4.4 – 7.0 mmol/L range) with only a slight increase in the percentage of BG < 4.0 mmol/L (1.4% vs 1.1%) for the simulation (Table 5). Results were not different if the lower guard rail was increased to 4.4 mmol/L (Case 2). Hence, the trade-off between workload, and clinical performance and safety, is clear in that a major reduction in BG measurements is achieved while safety and GC performance are maintained. The 4.4 mmol/L guard rail is also a better choice as a rolling window can be more effectively implemented, and safety is not reduced.

If the time between new interventions was allowed to vary up to 3 hours (Case 3) a treatment interval similar to that clinically observed (Table 5.5) is achieved, except that now most treatments are calculated from the CGM SG measurement. This trial shows the potential results of the integration of a CGM device into the STAR protocol without any further changes to the controller that currently utilises intermittent BG measurements in the same interval. It thus matches current intervention workload if there is no automation. The key metrics were all similar to the first sensor model virtual trial with 1 hour infusions (3.0 recalibrations per day vs 3.3, 87.0% BG measurements in the 4.4 – 8.0 mmol/L band vs 87.9%, 1.37% BG measurements < 4.0 mmol/L vs 1.37%, 0.014% BG measurements < 2.22 mmol/L vs 0.019%). This result suggests STAR can be safely used with the GlySure CGM SG measures, or any similar performing CGM sensor, without requiring hourly treatment updates, thus limiting or reducing the level of automation required, while significantly reducing workload.

Case 4 was not significantly different to Cases 1-3, and thus the added rolling window did not show any notable benefit in most of the key performance metrics. Safety was maintained, reflected in the percentage of patients who experienced a severe hypoglycaemic event (1.41% (10 patients of 708) vs 1.13% (8 of 708) for Case 1). Thus, rolling windows to capture rapid change impact safety relatively independent of performance and workload.

Figure 5.4(a) shows the percentage of BG measurements within the thresholds, and shows no change in percentage BG in band as reported by the sensor and no change in median BG (Table 5.7) for a given change in recalibration period. However, there is a decrease of percentage BG in band for the true BG as the recalibration period is increased. This outcome reflects the decreased accuracy of the CGM sensor as the sensor is allowed to operate for longer without recalibration, and particularly, the unseen impact of sensor drift [107-109].

Figure 5.4(b) shows no change in the incidence of severe hypoglycaemia as reported by the true BG or SG. Again, there is a slight disparity between true BG and SG, where the true BG has the incidence of mild hypoglycaemia increasing at a slightly higher rate. This disparity again reflects the decreased accuracy of the CGM sensor and unseen impact of sensor drift with longer recalibration intervals. Overall, results are robust to calibration interval with mild degradation over longer periods, as expected. Results are not worse, in these cases, due to the use of guard rails to prevent BG levels from "escaping" the desired range and "forcing" recalibration if required. Thus, it is only unseen drift when true BG moves but drifting SG remains in band that is an issue.

Figure 5.4(c) shows the inverse relationship between nurse workload and patient safety. A 6 hour period between recalibrations increased the number of recalibrations per day to 4.3, which is still a large reduction from current clinical workload, but did not significantly increase the performance or safety of the controller. As expected, there was a slight increase in BG < 2.22 mmol/L with increased time (12 hours) between recalibration measures. Overall, a maximum of 8 hours between recalibrations appears to be an appropriate balance of workload and safety with this level of sensor performance.

Surprisingly, the addition of a rolling window monitoring the rate of change of BG did not affect overall performance. This result could be due to the activation condition for the rolling window being in practice very similar to the condition for which the guard-rails activated a recalibration and intervention. The rolling window recalibration and intervention would activate earlier than the guard rail, and thus there would be no need for another recalibration if or when BG reached 4.4 mmol/L or 8.0 mmol/L if it occurred within an hour of the rolling window recalibration. Thus, the benefit of the rolling window is in capturing guard rail events earlier to improve safety.

Both subcutaneous and venous/arterial CGM devices have been studied for use in the ICU [79, 80, 95, 129-131, 136, 140, 172, 193-196]. Of these studies, Boom et al. [130] offers the most comparable study

to the one presented, where GC based on a sliding scale algorithm was guided by subcutaneous CGM measurements. Performance, as measured by the percentage of measurements in the target range, was worse (69% in the 5.0 – 9.0 mmol/L target range vs. 88.0% in the 4.4 – 8.0 mmol/L target range for Case 4 of CGM-enabled STAR) [130]. Safety was comparable between the two studies, with no incidences of severe hypoglycaemia (<2.22 mmol/L) in Boom et al. (0 out of 78 patients) compared to 1.41% of patients (10 out of 708 patients) in Case 4 of CGM-enabled STAR. Other studies in which GC was guided by subcutaneous CGM sensors reported average GC performance results alongside significant incidences of hypoglycaemia [79, 172], while another reported reduced incidences of hypoglycaemia but no improvement in GC performance compared with intermittent BG measures [129].

Other studies show a similar percentage time in target range, but are less comparable because they used blood gas analysers to measure BG and guide GC. In particular, several small studies show relatively high percentage of BG measurements in a tight range, with 0, or close to 0, incidence of severe hypoglycaemic events (BG < 2.22 mmol/L) [80, 95, 196]. Other studies have used blinded CGMs to assess CGM point and trend accuracy, through MARD and Clarke error grid analysis, or detect hypo- or hyper-glycaemic events that may not be measured with intermittent BG measurements, but were not used to guide GC [96, 115, 117, 168, 206]. Overall, this study uses clinical data from a large number of ICU patients to demonstrate the potential for using CGMs to aid or guide GC in the ICU, an outcome supported by preliminary results from previous studies.

The overall results of this study show replacement of intermittent BG measurements in GC may significantly reduce nurse workload, without compromising GC safety or performance. The simulated CGM sensor was able to decrease nurse workload, in the form of manual measurements, from 12.7 measures per day to 3.0-3.6 measures per day, depending on which guard rail or rolling window

definition is employed. This decrease in nurse workload is achievable, with no significant changes to control performance or patient safety.

This result is in line with expectations [173], which hold that more frequent measurement would aid control. Additionally, STAR already had a high level of performance [34]. Lesser performing GC protocols could see improvements not evident in this study.

5.4.2 Limitations

One of the limitations of this study was that in the making of the CGM sensor model, the clinical data used to construct the model had a very low number of BG measurements less than 5.0 mmol/L [107]. This limitation makes it harder to predict sensor behaviour at this lower BG region, and it was assumed that the sensor behaviour had consistent percentage error for the range of BG measured, implying smaller absolute errors at lower BG and higher absolute errors at higher BG. This assumption is consistent with other studies carried out using brands of interstitial CGM devices [98, 121, 184], while others report an increase in MARD at lower BG [114, 122, 185]. This limitation could have a sizeable effect on the data produced, as the percentage of BG measurements < 5.0 mmol/L recorded by the sensor ranged from between 10.1% to 13.8% for the sensor model virtual trials. Further research would need to target data in this BG range to ensure the reliability of the sensor model at lower BG. However, based on the initial validation the current model is likely acceptable. Similarly, if the model of the CGM sensor at lower BG is inaccurate, a rolling window in a clinical setting could prove more effective in catching rapid declines in patient condition and BG than the simulations suggest.

A further limitation is sensor signal delays. Both the signal filter and the diffusion through the sensor membrane will introduce some delay. However, with the relatively high frequency of sensor measurements (4 times per minute) and the relatively long period of time between treatment decisions for the STAR protocol (1-3 hours), this combined signal delay would not be a hugely significant source

of error clinically. Thus, it was not explicitly accounted for in the development of the model or the simulation of the CGM sensor traces.

5.5 Summary

A previously validated model of an IV CGM sensor was used in place of intermittent BG measurements to guide GC decisions under the STAR protocol in a virtual environment. Virtual trials are used to evaluate the impact on safety, performance, and workload of implementing these or similar CGM sensors for GC using the STAR protocol. Using a range of guard rails and/or rolling windows delineated trade-offs found in using CGM sensors, and decreased the number of required blood draws for BG measures by up to 73%, while also maintaining GC performance and patient safety. Performance and safety were robust to reasonable changes in sensor recalibration period, as well as to changes in rolling window parameters of window length and BG change. Overall, the use of a typical CGM sensor in clinically validated virtual trials shows the potential to reduce clinical workload significantly. The lack of equally significant improvements in performance are likely due to the already very good and clinically proven performance of the STAR GC protocol employed.

Chapter 6 Variability and Level Analysis of CGM Sensor Traces

6.1 Background

A recent study of metabolic variability and mortality outcomes showed no difference in underlying metabolic variability between survivors and non-survivors [32], suggesting patient outcomes are a function of the quality of control delivered and not patient-specific or cohort-specific characteristics. Poorly delivered control might thus affect study outcomes when looking for benefit or harm, as studies show good glycaemic control (GC) may need to be achieved for essentially all patients to achieve mortality reduction [48]. As such, there is a clear need for safe, effective control algorithms able to directly manage the significant intra- and inter- patient variability that makes GC difficult [49, 148, 207]. This outcome points to a clear need for greater input from the fields of control systems and automation to manage variability and provide control.

Model based methods have been developed to achieve this goal [34, 41, 54-56]. In particular, many monitor and respond to changes in patient-specific metabolic condition using an insulin sensitivity model parameter. Insulin sensitivity is a key determinant of the glucose uptake response to an insulin dose, and this sensitivity is most variable early in the intensive care unit (ICU) stay where most hypoglycaemia occurs [23], both in response to patient condition and clinical interventions [24, 57-59, 208].

However, while the future looks optimistic for model-based GC and patient outcomes [44], there is still much debate over what GC targets are appropriate and/or safe for an ICU context [39]. Many prefer higher targets out of fear of hypoglycaemia [26-29]. Some studies have shown greater benefit from lower targets [25, 26, 40, 64, 65]. Further, evidence suggests these targets may be patient specific and/or differ between ICU cohorts [66]. Confounding this issue is a lack of consensus on how to measure and/or report GC outcomes and variability at a cohort and patient level [30, 67], especially as some

model-based GC use variability [52, 53, 156, 157] to dose insulin [54, 156-158].

Thus, there is a clear need for coherency and novelty in regard to assessing glycaemia, both in level and variability, preferably in a fashion readily transferred to other clinical signals. Consensus in assessing state or level and variability with metrics well-correlated with clinically relevant outcomes would provide the means to assess all protocols and methods, where such metrics are lacking [30, 50]. Hence, while automation and control are emerging in this field, the key control parameter, the measure of desirable outcome parameters, glycaemia and its variability in this case, are missing or not well-defined.

This chapter summarises the state of the art with regards to assessing glycaemic level and variability in hospital ICU cohorts. It explores the common definitions, and their advantages or disadvantages. It then provides a vision for the future through a new, novel, state-based description of these quantities, in an attempt to better capture patient and cohort glycaemic behaviour. It is particularly applicable given the rise of CGM use in general, and in the ICU in particular [174, 209, 210]. The methods, review, outcomes and analyses are generalisable to a range of clinical metrics in critical care.

6.2 State and Variability: Measures of Metabolic Quantities

6.2.1 State and Level of Blood Glucose

Blood glucose (BG) concentration, or its average across a population or time period or both, is the most common metric reported. Assumptions around appropriate concentrations and how they should be achieved underpin much of the glycaemic control research and clinical practice. Common thresholds for blood glucose and dysglycaemia differ in derivation and definition between different intensive care cohorts, and between adults and neonates in particular.

6.2.1.1 Adults

Adult hyperglycaemia is well researched and BG greater than 8.0 mmol/L is the typical threshold defining hyperglycaemia [2, 35, 36, 54, 211-213]. This definition and the upper limit on GC target ranges varies within the literature from 6.1 – 9.0 mmol/L, where many individual clinicians, centres, and protocols define their own safety thresholds for normoglycaemia. One meta-analysis showed no particular advantage for GC ranges, although they found lower ranges were associated with higher incidence of hypoglycaemia [214]. However, several other studies have shown lower ranges, while avoiding hypoglycaemia, offer reduced risk of death [25, 26, 65]. These differing outcomes, using different ranges, are a confounding factor in the pursuit of both consensus and reduced clinical hyperglycaemia.

Thresholds for mild and severe hypoglycaemia in adults are also well defined [23, 28, 29, 54, 102, 215, 216]. The consensus is mild hypoglycaemia occurs at 4.0 – 4.4 mmol/L, while severe hypoglycaemia is most commonly defined as BG < 2.2 mmol/L [2, 130]. These thresholds result from clinical studies of decreased BG and the onset of physiological responses as a result of hypoglycaemia [217], and they may be patient-specific [218]. In particular, patients with type 2 diabetes and persistent hyperglycaemia can exhibit hypoglycaemic like symptoms at normal BG levels [218].

6.2.1.2 Neonates

In contrast to the well reported clear limits of normoglycaemia for adults, normoglycaemia for infants is less intensively examined and defined. As a result, clinical practice and research definitions differ. Alsweiler et al. [219] surveyed 27 tertiary neonatal units in Australasia for their definition and management of hyperglycaemia in very low birthweight infants. Variance in the clinical definitions of hyperglycaemia from these specific neonatal units ranged from 7.0 – 15.0 mmol/L, though most used 10.0 mmol/L.

Alsweiler et al. also observed large variance in target ranges for subsequent insulin therapy, though many defaulted to the normal ranges for adult glycaemia of 4.0 – 8.0 mmol/L. In some cases, the upper target ranged between 8.0 – 10.0 mmol/L. Some of the lower ends of these target ranges were much lower at 2.5-3.0 mmol/L, reflecting evidence of a possible tolerance for lower BG by infants [220, 221]. Thus there is overall little consensus for onset criteria and targets in neonatal GC, reducing effectiveness of protocols, particularly with respect to later outcomes, and ability to compare study outcomes.

In contrast to adult definitions, and also to the clinical thresholds observed by Alsweiler et al., the thresholds for neonatal hyperglycaemia used in literature by researchers is significantly lower, from 6.9-8.3 mmol/L [222-226]. They are based on observed distributions of BG in preterm infants [227], where less common extremes are used to define abnormal BG. However, these statistical definitions may simply describe measured BG in this cohort, rather than what may be clinically desirable, beneficial, or obtainable via GC [227].

Neonatal hypoglycaemia is controversial, where agreement has not been reached on either thresholds or treatment [220, 228-230], in part because it is often asymptomatic. Some studies show neonatal BG concentrations reach natural lows at approximately 2 hours after birth [231], while others fail to observe this nadir [232, 233]. Infants experiencing hypoglycaemia may also be asymptomatic [234-237], unlike adults, and it is still unknown if asymptomatic hypoglycaemia is the same as symptomatic hypoglycaemia, or if such thresholds are patient-specific. However, neonatal hypoglycaemia (defined as <2.6 mmol/L) has been linked with poor neurologic outcomes in later life [221], and thus better definition in this area would lead to improved treatments and outcomes.

6.2.2 Measures of Glycaemic Variability

Though perhaps intuitive qualitatively, glycaemic variability is difficult to effectively quantify. Many reviews have undertaken to summarise variously measures of glycaemic variability, particularly in the

context of Diabetes management, and more fully describe their advantages and limitations [75-78]. Many metrics, and adaptations of metrics, exist, but in general most fall into the following broad categories: 1) descriptors of middle and range; 2) descriptors of total or summed variability or excursion length; and 3) descriptors of time in range.

The most standard methods for describing and quantifying glycaemic variability are statistical descriptors of middle and range in data. The most common descriptor of variability is to report the standard deviation (SD) of BG measurements alongside the mean BG value [79]. Similarly, the IQR (inter-quartile range) is a non-parametric alternative, reported alongside median BG. These metrics are popular because of their ease of use [75]. However, the SD is a measure of dispersion, rather than variability, and is limited in its ability to reflect the time course of BG measurements [80].

The limitations of the SD for describing glycaemic variability are shown in Figure 6.1, where the mean and SD are the same for two time courses of intermittent BG measurements. In this case, the same set of BG measurements is arranged in two different patterns, one a decrease at a constant rate, and the other moves between extremes. Because both have exactly the same measurement set, both have the same mean and SD. If glycaemic variability is more truly a function of the change in BG, with more rapid changes exerting different physiological effects than slower changes, than patient outcomes could differ in these two cases. Thus, the SD is extremely limited in its ability to describe the time course of glycaemic variability.

In addition, the mean and SD assume a normal distribution, which is inaccurate, particularly in dysglycaemia, as BG is usually highly lognormal and skewed. Thus, this measure for variability may be inappropriate to use, particularly if there are a large number of measurements near the hypoglycaemic or hyperglycaemic ranges of BG. Equally, all such measures requiring a normal distribution assumption have this issue, and recent reviews recommend non-parametric statistics [30].

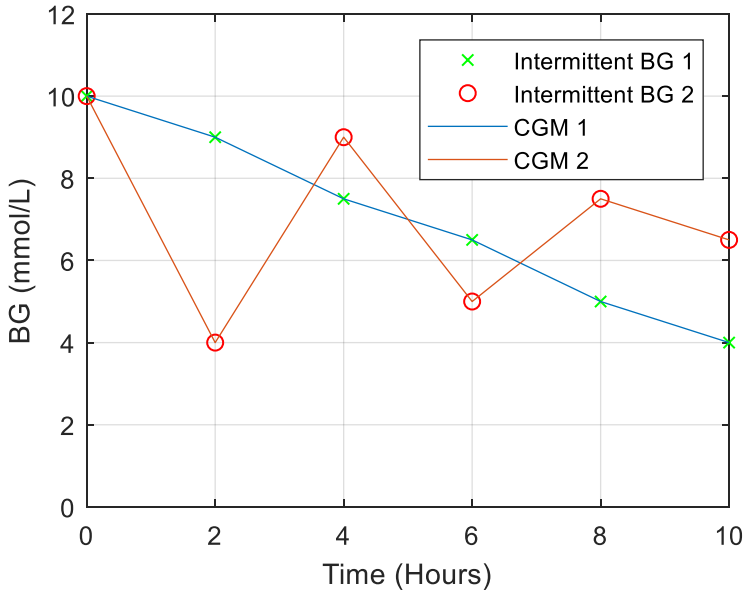


Figure 6.1: Two sensor traces with accompanying intermittent BG measures. The intermittent BG measures are the same between the two data sets, and so have the same mean and standard deviations, but display differing behaviours.

The Coefficient of Variability (CoV), which is the mean divided by the standard deviation, is for some a preferred, more intuitive, measure of glycaemic variability [75]. CoV can be a good measure of overall glycaemic variability, as research has shown that aiming for a CoV lower than a certain threshold (36%) allows for the distinction between stable and unstable glycaemia [81, 82]. However, this metric suffers the same limitations as discussed regarding the mean and SD above. For the data in Figure 6.2, the CoV would by definition be the same between these two patients.

Metrics examining total glucose excursion include Area under the BG curve (AUC) and Glucose Miles. AUC or area around a line is common in diabetes and non-critically ill cohort studies, often with more frequent continuous glucose monitoring (CGM) [83-85]. Two zig-zagging sensor traces could theoretically have the same area, even though one trace is rising in level, while the other is falling. This behaviour would suggest the two traces have the same level of variability, but with differing behaviours, indicating how AUC is an overall measure of variance, but not specific to time course. This issue is also illustrated in Figure 6.2.

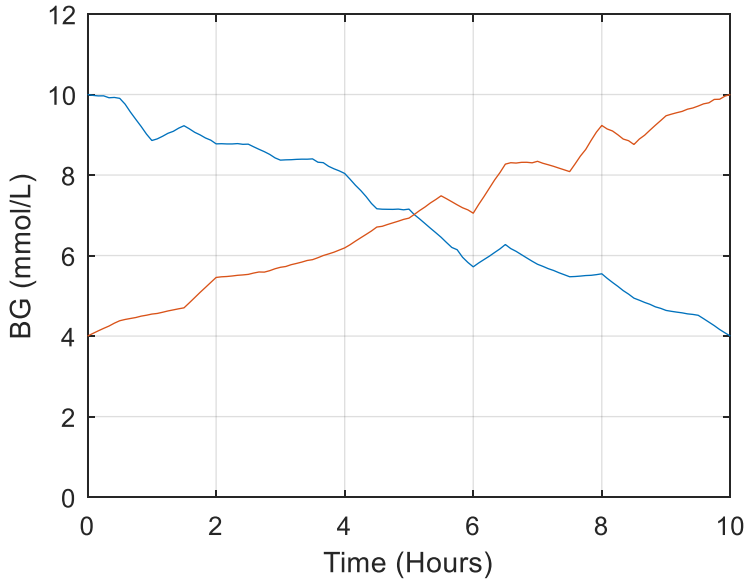


Figure 6.2: Two sensor traces with very similar AUC, but differing behaviours.

Glucose Miles is another way to measure variability, measuring the total 'distance' travelled by movements in BG through intermittent BG or CGM traces. It is embedded as part of some other measures [86, 87]. However, two traces can have the same Glucose Miles with very different mean BG, as shown in Figure 6.3 for a bias, and Figure 6.2 where the median is the same. Thus, this metric is similar to AUC in giving a total, but is not specific or reflective of the time-varying behaviour of blood glucose.

Furthermore, while Glucose Miles and AUC can be useful descriptors of cohort variability when paired with measurements of mean or median BG, both have an inability to describe variation away from some longer term average or glycaemic state, which may be important to recognise clinically. Similar metrics, more suited to intermittent BG measurements, include the mean absolute difference (MAD), or mean of daily differences (MODD). Such metrics describe the average change in BG, or the difference in BG between days, but are limited in their ability to comprehensively describe the time course of variability and its relationship to the level, as per Figure 6.3.

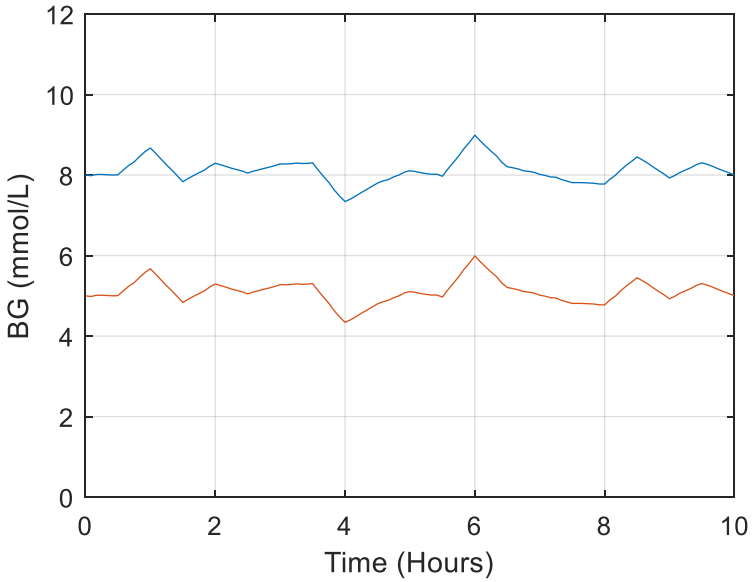


Figure 6.3: Two sensor traces with similar Glucose Miles, but with very different mean BG.

Time in band is another typical measure of variance good for combining and capturing level and variability, though it tends to be more of an implicit measurement, rather than explicit. It is a good aggregate measure for larger cohorts where central tendency makes time in band informative of overall cohort behaviour. Research has shown time in band, or time in range, can be associated with clinical outcomes [25, 26, 64, 65], in particular risk of microvascular complications [91].

However, whether the measurement is inside or outside of the band itself is discrete or binary. Thus, all variability or measures in the band are assumed to be clinically acceptable, and those outside are not. It may be a measure just inside the band is thus treated very different in analysis to one just outside, when both could be within measurement error. In addition, there is no agreement on appropriate bands, leading to difficulty in comparing the variability across studies [30]. By example, Figure 6.4 shows two sensor traces for a 4.4 mmol/L to 8.0 mmol/L range, which have the same time in band, but different, potentially clinically important behaviours both within and outside the band.

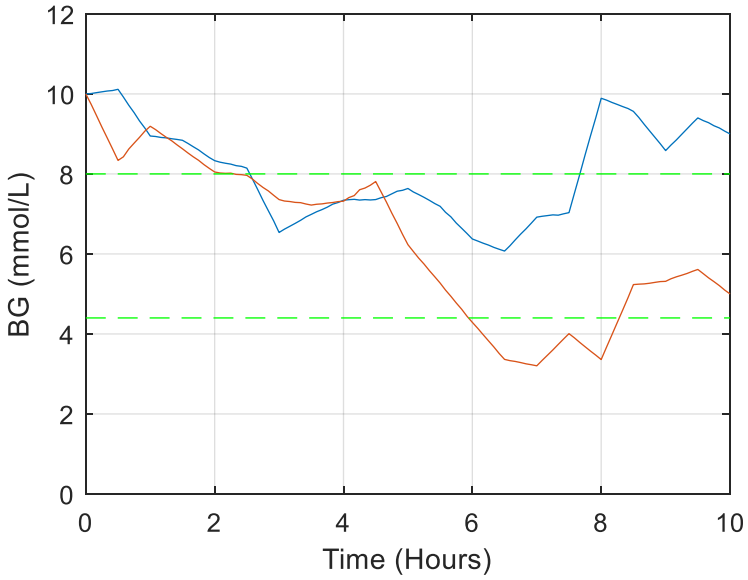


Figure 6.4: Two sensor traces with similar time in band, but differing clinical behaviour and impact.

Other metrics for describing glycaemic variability exist, such as the mean amplitude of glucose excursions (MAGE), M-value, J-index, Low Blood Glucose Index (LBGI), High Blood Glucose Index (HBGI) and Average Daily Risk Range (ADRR) [75-78, 92]. Each of these metrics attempts to compensate for some aspect of the limitations of the metrics previously discussed, and specifically attempt to account for aspects of desirable glucose outcomes. Most of these metrics involves one or more aspects of the middle-range, time in band, or excursion size categories. All require some degree of correlation with outcomes to be clinically applicable, and have shown some degree of predictive capability [92]. While their disadvantages are presented elsewhere [75-78], the overall common disadvantage of such metrics is their calculation and interpretation is not transparent or intuitive to those unfamiliar with the metric, which may have impeded more widespread clinical uptake.

Furthermore, most of these glycaemic variability metrics have been derived using point glucose measurements, and most have only been validated using these intermittent BG measurements to check the correlation with adverse outcomes. Only one of the metrics has been validated using CGM sensors, and was found to be correlated with increased hypoglycaemia given the measure of glycaemic

variability over some threshold [82]. However, more studies are needed for CGM sensor derived metrics and correlation with adverse outcomes.

In addition to the various existing metrics for variability, variability in itself has been hard to define depending on what is clinically relevant. For example, observing the two sensor traces in Figure 6.1, on the timescale presented of 10 hours, the two sensor traces clearly have two differing levels of variability. CGM sensor trace 1 is far less variable than CGM sensor trace 2. However, stretching the timescale to a 24 hour period, the two sensor traces could be interpreted as having similar variability. Thus, there are currently no metrics that can adapt their timescales to account for the full time course information CGM sensors can provide.

In summary, the large range of glucose metric definitions are confusing, contentious, and complicated. No metric adequately describes the time-course of glycaemia. Variability, in particular, is poorly captured by commonly used metrics, most of which are mathematically, rather than clinically, centred and defined. These limitations are particularly important in the context of emerging CGM technologies, which provide greater time resolution with greater, increasingly enhanced, point accuracy compared to intermittent measures [93, 115, 119, 238, 239]. CGMs thus offer the potential to significantly improve GC. However, to obtain this benefit, the best clinically defined metrics to capture patient glycaemic level and state must be established, which may require a new approach to defining variability. One that is clinically focused, rather than mathematically focused.

6.3 CGM Technology

CGM devices show significant as yet untapped potential for improving glycaemic monitoring and control, particularly due to their high measurement frequency. CGM sensors measure BG near continuously, with new generation devices able to measure at a rate of 1-6 times per minute [93-96],

and standard devices providing measurements every 5 minutes [97-100]. Compared to 1-6 hourly point-of-care (POC) measurements in a well-staffed ICU [101], these devices offer huge potential to improve care and reduce workload in adult [174] and neonatal ICUs (NICU's) [240].

The increased measurement frequency has many benefits for care, including the ability to monitor patient condition and, importantly, the trajectory of their glycaemia and condition in real time [167, 196]. They also provide warning for hypoglycaemic events [87, 102], allowing early correction. Both benefits cannot be achieved with intermittent BG measures, where the minimum feasible regular measurement interval is ~1 hour, but clinical non-compliance can be high even at lesser rates [103-105], as seen in a recent study where protocolised measurement interval was 1 hour and the clinical measurement was closer to 3 hours [31, 39]. Even then, a patient's condition may change significantly between hourly POC BG measures, resulting in hypoglycaemia remaining untreated for up to 50 minutes. A CGM sensor, on the other hand, can alarm both at occurrence and predictively before it occurs.

Despite these advantages, CGM sensors are not widely used in the ICU as CGM sensor technology still suffers limitations, including larger point error inaccuracies and sensor drift [107-114]. They are also expensive, so it may be hard to justify the cost versus potential benefits. The larger point accuracy errors over traditional intermittent BG measurement techniques have been well documented, with new CGM devices usually reporting gradually improving MARD values up to 8-12% [93, 115-117] as the technology has developed [93, 95, 98, 109, 115-125], compared to the 5% or lower common in intermittent POC measures [126]. Sensor drift is also still not widely recognised, even though it has shown to be a key driver of larger MARD and potential hypoglycaemia when used in GC [107, 127].

CGM technology also has the potential to reduce GC related workload in the ICU, providing more bedside data with lower blood sampling requirements [127-130]. Clinical practices have cited high

workload as a reason for reduced intermittent BG measures, as not every ICU has a low staff to patient ratio to justify the increased workload some GC protocols may require [68, 70, 71]. CGM sensors are able to give a continuous readout of the current BG measurement, and clinicians can adjust nutrition and insulin accordingly. However, this benefit has not yet reached regular care due to the limitations.

Overall, CGM technology has not reached regular clinical use in ICU care, despite common acceptance in lower-risk outpatient type 1 diabetes care [159-163]. As CGM technology improves it will be able to replace intermittent measures used in the ICU today. However, with increased temporal measures and benefits to control, the same issues of how to quantify level and state remain. In fact, with the increased measurement rate of CGM technology, these issues are exacerbated as measures that work for intermittent measures may not be representative or accurate for use with CGMs. There thus remains a significant need for better, more clinically defined metrics, and in particular, those able to maximise the measurements delivered by CGMs.

6.4 Vision for Glycaemic Control Metrics in the Future

Given the development of increasingly accurate CGM devices, which offer higher resolution glycaemic monitoring, and the limitations of current measures of glycaemic level and variability, improved methods must be developed to describe glycaemia. Such methods will be important for effective research into cohort and patient specific behaviours and the benefits and harms of dysglycaemia and insulin therapy. These metrics should be clinically defined and thus associated with clinical outcomes, avoid reliance on statistical assumptions, easy to implement computationally in real-time, and simple to understand to aid uptake.

Our vision for the future is a method for describing the time-course of glycaemic level and variability. Such a method would both improve reporting of glycaemic outcomes and better aid bedside glycaemic monitoring and treatment in patients. The measure would be able to fully utilise the increased resolution

of modern CGM sensors allowing for analyses to relate the glycaemic states, and critically, any changes to these states, to clinical outcomes. The measure would be readily generalisable to other aspects of measurement and care, and would be easily compared across centres and cohorts under glycaemic control.

6.5 Glycaemic State Analysis

This section presents a novel method with the potential to fulfil the vision for the future defined. It develops a characterisation method for identifying Glycaemic States utilising CGM sensor data to monitor patient condition, both retrospectively, and in real-time. The algorithm builds on the idea of patient specificity, and also provides scope to adapt to changes in patient metabolism or clinical goals. The result is demonstrated on a cohort of 614 infants at risk for hypoglycaemia, and provides a first CGM-centric approach to simply quantifying Glycaemic State and variability in real-time. Infants were used for this analysis, as the data was conveniently available, and it was perfectly suited for the application being glucose dense CGM sensor data. It also had the additional benefit of being a study that explicitly looked for changes in glucose state around hypoglycaemia and normoglycaemia, and their relationship with outcomes.

6.5.1 Using CGM Sensors to Understand the Evolution of Glycaemia over Time

During the CHYLD study, the CHYLD researchers recognised qualitatively that the infants seemed to transition through distinctive Glycaemic States. These States, described by the clinicians as periods of relatively stable glucose levels over a period of 3-6 hours or longer, were patient specific, and seemed to vary not only from patient to patient, but also varied a lot within a patient. The research group thought there might be links between these Glycaemic States, and the neurodevelopmental outcomes of these infants presented in earlier studies [220, 221].

However, the research group had no way to quantify this variability or hypoglycaemia at the patient specific level, instead relying on visual inspection of the CGM sensor trace. Thus, this work set out to

design such an algorithmic tool to automatically quantify these Glycaemic States from these CGM sensor traces. This work would involve consultation with the clinicians in the CHYLD research group to establish what constitutes a State or State Change, and how these State Changes might be related to the clinical diagnoses of poor neurodevelopmental outcomes.

In this chapter the ideas of lengths of a State and size of State Changes are explored, based on information gathered from clinicians to quantify behaviour that is of interest. The parameters are then set on clinician recommendations in order to use the algorithm on the CHYLD patient CGM sensor traces. Later in Chapter 7, these parameters are varied and tested against the associated neurodevelopmental outcomes to evaluate the effectiveness of this tool in aiding diagnosis. This chapter thus presents a first step towards a patient-specific description of state and variability.

6.5.2 Calculating Patient States

For each CGM trace, the mean of the whole CGM trace establishes a baseline average interstitial glucose (IG) level for that patient. The CGM data is then filtered using a centred 6-hour rolling average. Thus, there are 12x6-hour-average data points created using every hour of CGM data available (CGM sensor measurement frequency = 5 minutes), commencing 3 hours into the trace as the rolling average is calculated from the centre of the rolling window.

The 6-hour rolling average is thus bootstrapped by taking the average of the first 3 hours of the CGM trace to begin the rolling average. The rolling average is then rolled forward by one measurement from this point for the next 3 hours until it reaches 6 hours. This method is reversed for the end of the CGM trace.

Comparing the 6-hour rolling average to the arithmetic mean of the entire CGM trace yields a baseline variation around the overall mean value. These variations are examined for size to find state changes. Six hours is chosen because it filters out higher frequency glucose fluctuations, such as spikes just after feeding in the NICU, and allows for any true, long term changes in average IG, characterised as a State

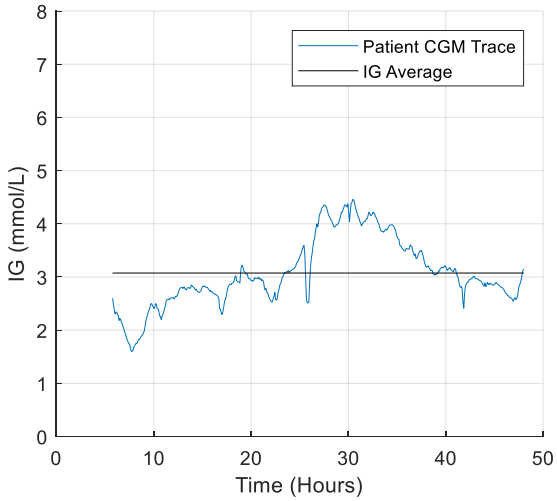
Change in patient metabolic behaviour.

Each time the 6-hour rolling average crosses the arithmetic average line, it is considered a possible State Change if specific clinically defined conditions are met:

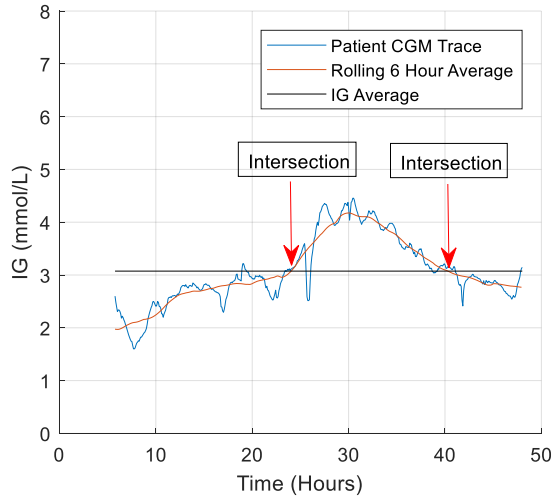
- More than 5 hours have passed since the last State Change, which assures "States" are 5 hours (or longer) periods of relatively constant average IG.
- The State Glycaemic Average, defined as the average IG for a given State, was more than 0.3 mmol/L higher or lower than the previous State Glycaemic Average, to reduce potential impacts of measurement error.

If either condition is not met, the CGM data for the current arithmetic average crossing is combined with the CGM data of the previous State, and that State's Glycaemic Average recalculated using the longer CGM data. This State characterisation process is shown in Figure 6.5.

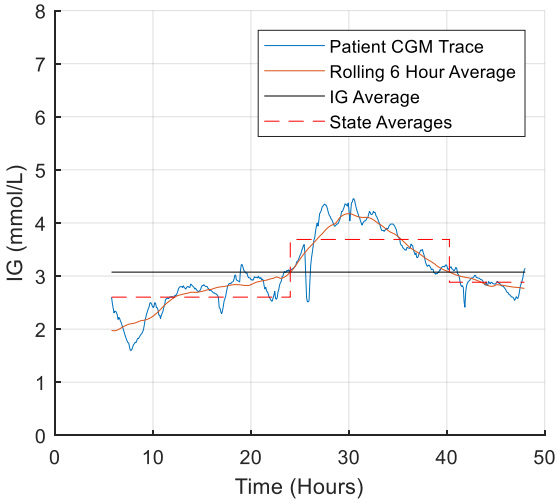
Importantly, while the 0.3 mmol/L threshold used here is based on clinical experience with neonates, it can be changed for any cohort or to find only larger changes. Similarly, a shorter or larger rolling average than 6-hours can be used, or they may be "nested" to find shorter State periods. Finally, State Changes per day (24 hours) are calculated to normalise results for comparison, but, equally, they could be calculated in a roll-forward manner for use in real-time.



a) Clinical data for Patient 1. Black line shows the average for all of the CGM data.



b) A centred rolling 6-hour average is calculated. Intersections between the rolling 6-hour average and IG average are noted by the red arrows. The rolling average is also bootstrapped to fill in the 3 hour gaps at the start and end of the rolling average line.



c) The average for the Glycaemic States are calculated and plotted.

Figure 6.5: Steps of the Glycaemic State characterisation process, showing how the rolling average and States are calculated.

The median [IQR] hours of CGM per patient are calculated, alongside the average change in sensor measured interstitial glucose (IG) after a State Change, the maximum State Change, the minimum State Change, number of State Changes that resulted in a higher IG (also as a percentage), and the number of State Changes that resulted in a lower IG (also as a percentage). The median IG average is also calculated, alongside the IQR and 90% range, to allow for comparison of this cohort to others. This is a search, given the method as defined by the CHYLD clinicians and statistician, for links between IG states and state changes, with outcomes. In particular, babies with more variable CGM traces were identified, qualitatively, by CHYLD researchers and clinicians as having poor outcomes. Thus, we are seeking to quantify that link.

6.5.3 Case Study from a Neonatal Cohort

6.5.3.1 Subjects and Continuous Glucose Monitoring

The CHYLD Study recruited 614 infants born from 32 weeks gestation with one or more risk factors for neonatal hypoglycaemia, including the following: diabetic mother; preterm (<37 weeks); small (<10th centile or <2500 g); large (>90th centile or >4500 g); or acute illness [220]. The aim was to examine the relationship between the incidence and severity of neonatal hypoglycaemia in at-risk infants and neurodevelopmental outcome in childhood [220, 229]. A total of 481 infants had an interstitial CGMS System Gold CGM sensor (Medtronic Inc., Northridge, CA) inserted soon after birth in the lateral thigh, as previously described [229, 241]. Of these infants, 366 had more than 24 hours of CGM sensor data in the first 48 hours after birth, leading to 12356 total hours of CGM data (median [IQR]: 35.7 [30.5 38.4] hours/patient). Twelve further infants had more than 24 hours of CGM sensor data in the first 48 hours after birth, but had gaps in the data of more than 5 minutes (mean [IQR] = 8.1 [3.0 8.9] hours) and were thus excluded from the analysis. The CGM sensor recorded a measurement every 5 minutes, but results were masked and did not influence clinical care. CGM data were downloaded and recalibrated to all blood glucose concentrations, measured on a blood gas analyser [242]. The study was approved by the New Zealand Northern Y Ethics Committee.

6.5.3.2 Results from a State and Variability Analysis

Table 6.1 presents the overall results of the State analysis algorithm. The number of State Changes experienced by each infant was calculated, along with the number of State Changes/day, average absolute change in IG over a State Change, minimum and maximum overall State Changes over the entire cohort, and the number of State Changes with higher and lower IG. The results reported here give greater depth to glycaemia, highlighting the frequency of glycaemically stable or variable patients.

Table 6.1: State Change analysis results.

Patients	366
Total hours	12356
Hours/Patient (median) [IQR]	35.7 [30.5 38.4]
Median IG Average [IQR], (90% range) (mmol/L)	3.7 [3.3 4.1], (3.0-5.0)
Number patients with no State Changes	56 (15.3%)
1 State Change	177 (48.4%)
2 State Changes	81 (22.1%)
3 State Changes	46 (12.6%)
4 State Changes	5 (1.4%)
5 State Changes	1 (0.3%)
State Changes/day (median [IQR], (90% range))	0.65 [0.54 1.16], (0-1.73)
Median [IQR] absolute ΔIG State Change (mmol/L)	0.65 [0.45 0.92]
Max State Change (mmol/L)	3.7
Minimum State Change (mmol/L)	0.3
Number of State Changes from lower to higher average IG	311
Number of State Changes from higher to lower average IG	191

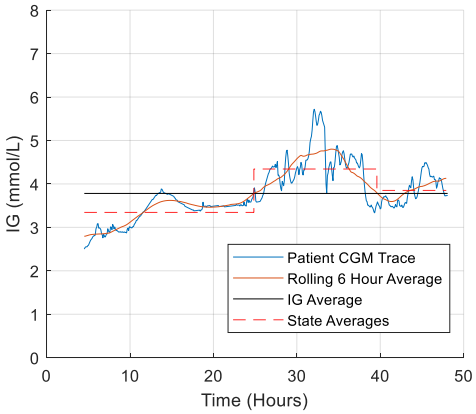
Figure 6.6 presents case studies from the State characterisation process. Figure 6.6(a) is a patient with 3 separate States identified. The initial patient State is low IG and rising, shown by the State Average of ~3.5 mmol/L. At around 25 hours after birth, the patient enters a State of higher average IG, but also higher variability, as shown by the large fluctuations in the original CGM signal about the State Average of ~4.4 mmol/L. At 38 hours after birth, the patient enters a steep decline in IG. However, this decline is not reflected in the rolling 6-hour average as a State Change until 40 hours after birth, where the rolling average intersects the IG average line. This State is less variable than the previous State and also has an average closer to the total IG average, implying the patient becomes more stable.

In Figure 6.6(b) only 2 States are identified, despite the CGM trace intersecting the IG average line

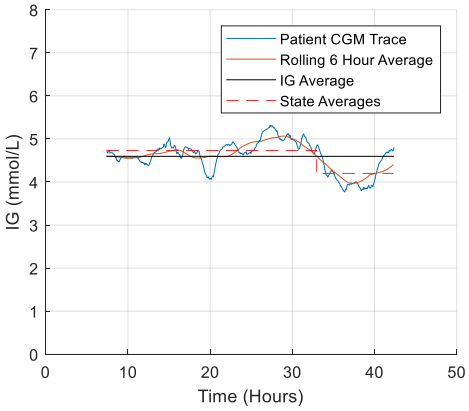
frequently. This is because the resulting States are less than the 0.3 mmol/L threshold required to reach a new Glycaemic State, and so these States are merged into one State. At approximately 33 hours, both the CGM trace and 6-hour rolling average decline enough, and for longer than the minimum 5 hour time frame, to describe a new, lower Glycaemic State. This patient has thus displayed a consistent and prolonged drop in glycaemia, which could be clinically significant.

Figure 6.6(c) shows a patient with similar States and State Changes to Patient 2 in Figure 6.6(a). The patient starts at a lower State, experiences a State Change to a higher level, and then drops again to a State Average closer to the overall IG average. In this case, high variability results in the rolling average crossing the IG average more frequently, although subsequent changes would not be significant (<0.3 mmol/L threshold). Clinically, it is of particular note that the IG spike at 37.4 hours after birth from 3.4 mmol/L to 7.7 mmol/L occurs in 20 minutes of measurements in the absence of parenteral dextrose boluses or buccal dextrose gel, and is likely to be sensor error, rather than a rise in true IG [243], which the clinically defined approach discounts because the clinical definition of a State is far longer. In contrast, every other mathematically defined metric would be skewed by this error.

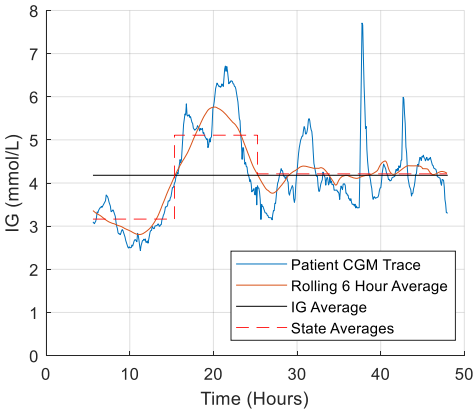
Figure 6.6(d) shows a relatively stable patient with regard to State, similar to Patient 3 in Figure 6.6(b), but with higher variability around that State. If a threshold >0.3 mmol/L was used, the algorithm would identify only one State. Thus, the choice of threshold depends on the granularity desired in defining clinically significant States. It is possible to change the algorithm threshold to determine an optimum value for State differentiation based on clinical outcomes, interventions, and measurement error thresholds.



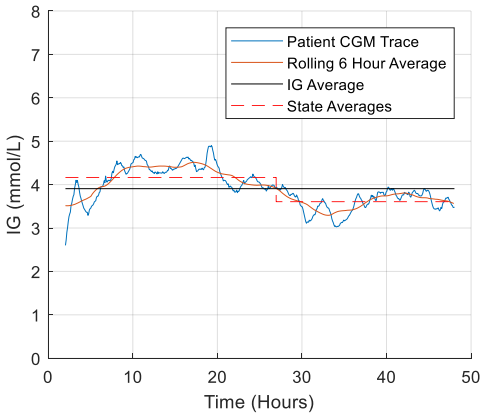
(a). State characterisation for Patient 2.



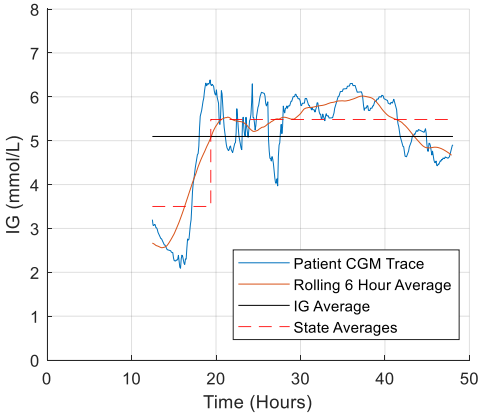
(b). State characterisation for Patient 3.



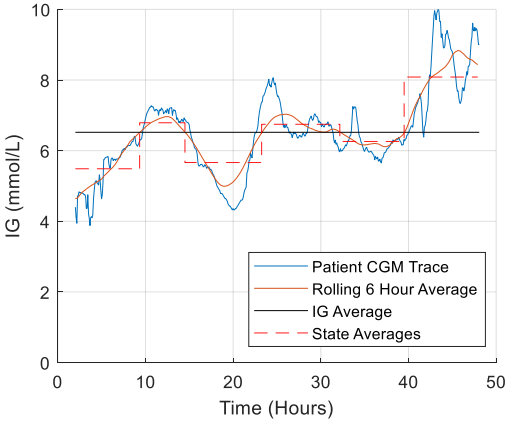
(c). State characterisation for Patient 4.



(d). State characterisation for Patient 5.



(e). State characterisation for Patient 6.



(f). State characterisation for Patient 7.

Figure 6.6: CGM data and Glycaemic State characterisation for Patients 2-7.

Finally, Figure 6.6(e) shows the State characterisation for an infant characterised by the algorithm as having 2 States, with moderate variability of the CGM trace. At the start of the CGM trace, IG starts low and appears to be hypoglycaemic for around 2.5 hours, appearing to be at a single State of low

glucose concentrations. IG starts to rise back to normal levels at around 17 hours after birth. The 6-hour average crosses the IG average again at 43.5 hours, but does not exist long enough to define another State. Thus, only 2 States are characterised. However, different threshold choices or a change in beginning the moving average could capture such early potential States, if clinically relevant and desired, showing how the clinical definitions lead this algorithm and metric, thus better relating the metric and clinical goals.

Only 1 patient recorded 5 State Changes in the 48 hour period, the maximum number for this cohort with the State definitions used. This variability could have been due to the infant being highly sensitive to feeding, where in the original data, the infant was fed at 4, 21.5, 36 and 42 hours, which roughly coincides with the large increases in IG. Figure 6.6(f) shows this patient, where it is clearer the IG average is abnormally high for this cohort. In particular, in the last portion of the CGM trace the infant is hyperglycaemic.

6.5.3.3 Discussion of Case Study Results

The majority of infants experienced less than 2 State Changes in the first 48 hours of birth (233 of 366 patients, 64%), suggesting these patients at risk of hypoglycaemia remained in a relatively stable condition despite high rates of hypoglycaemia [229, 241]. There were more State Changes from a lower IG to a higher IG than the other way around. This bias may relate to fact the CHYLD cohort recruited infants at risk of hypoglycaemia. Equally, BG tends to rise in days after birth as the infant's metabolism stabilises and their ability to take up nutrition and absorb it develops.

Only ~25% of infants (96 out of 366) were parenterally fed. The majority were entirely enterally fed. Thus, these State Changes are most likely not due to changes in parenteral feed rate, an outcome negated by the clinical choices of how a State is defined in size and length of time. The States, defined by the 0.3 mmol/L threshold chosen here, are by design larger than the expected rise due to a single enteral

feed. The 5-hour minimum State length is greater than a feed interval. Thus, it is unlikely that feeds would cause false State Changes as defined and used here, which again relates the metric definition to the specific clinical situation and goals for the cohort and is not possible with other current metrics.

With respect to variability, it is captured first by State Changes per day, as discussed. It can also be further enhanced by examining area around the State level, a local AUC to compare or assess variability within a State. Similarly, a second level State analysis could be run with different clinical definitions to capture Sub-States. The method presented is thus clinically defined, as well as being very flexible in this regard.

More adult CGM data is needed to compare to the infant results to see if there are any difference to level and variability between these two cohorts, or how much definitions of levels and States might need to vary between these cohorts. The method is generalisable enough to work with any CGM sensors and data, so using the method on other cohorts' CGM data would yield valuable results to compare with this infant cohort or others. Equally, there is room to define the metric specific to clinical goals and cohorts, or even to compute multiple metrics simultaneously relevant to different clinical outcomes within a cohort.

6.6 Discussion

6.6.1 Comparison to Glycaemic Metrics

The method of characterising States and State Changes offers some advantages over the mean and standard deviation, median and IQR, CoV, AUC, Glucose Miles and time in band metrics. This method for level and variability characterisation allows for increased resolution and descriptiveness in the reporting of individual and cohort glycaemia. In particular, stability to a particular 'State', and the variability around that State, are easily reported and intuitively interpreted. A particular strength of the method is in its ability to capture any changes, as critically ill patients can experience rapid changes in metabolic dynamics [106, 244].

In comparison to mean and SD alone, or CoV, this method offers far greater insight to patient condition. As shown in Figure 6.2, two BG data sets can have the same mean and SD but very different glycaemic dynamics. A well-tuned state and variability analysis would distinguish between these States. Importantly, this method can distinguish between variability around a State, and variability between States, both of which have different time constants and thus may have different clinical implications. Similar outcomes hold regarding AUC and Glucose Miles.

The metric used in this study is considered clinically relevant because it is clinically defined, where most prior metrics are defined mathematically, applied to clinical data, and then associated with a clinical outcome of interest to show its clinical relevance. In this case, the methods presented allow clinicians to derive the variability that is relevant for the issues/situation they are examining, while still assessing it in a uniform mathematical fashion.

6.6.2 Potential Clinical Uses

As noted, potential clinical uses are widespread. Foremost, is real-time use to allow visual inspection of changing glycaemia and variability around a State as it emerges and thus to track patient state. Such analysis could lead to better inputs to predictive models around patient outcomes [245-249].

Research uses are the likely first choice, where they would enable better analysis between and across studies. The simple calculations mean any data set could be calibrated to local clinical norms for comparison or to a standard set of definitions. In both cases, better comparison would lead to better understanding of the success and failure of GC interventions, thus enabling optimised care.

6.6.3 Methodological Strengths and Limitations

An important strength of the method is that all aspects are clinically defined. Thus, the parameters can

be changed in time or size to reflect the clinical behaviours to be captured. Hence, it is clinically defined and not dependent on any statistical distribution or assumptions. It can also be tuned to reflect clinical outcomes, where State thresholds may be adapted to reflect physiologically important changes in glycaemia.

The method does have some disadvantages. Currently, the method has only been used with retrospective data, and its use in real time and on other neonatal or adult cohorts has not been thoroughly analysed. Real time processing and updated States can be achieved with filtering and bootstrapping methods, and expansions to other cohorts simply requires the application of the method to available data. The part of the method where the whole sensor trace average is calculated and set as a baseline would also need to change, as the whole sensor trace would not be available in real time. However, the method can change slightly to be based off a moving whole IG average that continually updates as more data is received from the sensor. Although this may have room for some erratic behaviour at the start of data collection, the IG average should stabilise as more and more data is collected.

Another limitation of the method is that a new State can only be identified if the roiling average of the CGM sensor trace crosses the average glucose, IG. If, for example, a patient was highly variable in the short term above their IG, and then highly variable below their IG, only 2 States would be identified. This is partially affected by what the clinical parameters are set as for the method, for example setting shorter periods of interest can identify these shorter periods of variability, if that is what is of interest clinically. The variability within States can also be measured by combining with one of the metrics previously mentioned, such as Glucose Miles or AUC, if deemed appropriate for the analysis to be undertaken.

The method has not been applied to a data set with clinical outcomes to distinguish what State durations or thresholds are clinically significant. This future work will validate the overall usefulness of this method, and its contribution to the vision of improved descriptors of Glycaemic State and variability. Future work would also look to improve the method by making it robust to gaps in sensor readings

using interpolation or other methods.

Reported outcomes are still a simplification of the time-course of states and variability. While the State trace can easily and intuitively show changes in states and variability, attempts to reduce its properties to reported numbers in results or tables will inevitably result in some loss of information and intuitive interpretation of results. However, this issue is always going to be the limitation of any method to describe variability as a summary metric over any time period, and this method has the potential to allow greater description of glycaemic results in tabular or numerical form. For these results in particular, CGM sensor traces with length between 24 hours and 48 hours long have been pooled together in Table 6.1, which may hinder easy interpretation of these results. Thus, the data for the median and IQR State Changes per day were included to mitigate this increase in difficulty in interpretation.

6.7 Summary

The vision, analysis and review presented has addressed an emerging and critical aspect of glycaemic control, the need for consensus summary metrics of glycaemic level or State and variability. Current metrics have significant limitations as most are not clinically defined, and thus poorly represent many aspects needed for a control metric. It thus presents and defines the problem, current state of the art, and presents a vision for the future of what is required, including a first research effort to meet the goals defined. The overall work includes an ongoing focus to relate the ideas and current state back to traditional control systems methods and engineering approaches. There is still significant room for innovation and new approaches, particularly as more and more data becomes available to engineers, and clinicians become increasingly willing to take on more novel approaches driven by rapid changes in technology and the ability to measure and control patients. These changes will only occur more rapidly, driven in major part by the increasingly poor economics of healthcare mixed with the need to keep care more affordable via increasing automation.

Chapter 7 Variability and Level, and its Relationship with Clinical Outcomes

7.1 Background

As discussed in Chapter 6, normoglycaemia in infants is far less well studied than in adults. The literature thresholds are generally lower for both neonatal hyperglycaemia [222-226], and neonatal hypoglycaemia [220, 221, 250]. However, for neonatal hypoglycaemia in particular, agreement has not been reached by consensus on exactly what the thresholds are or what the treatments should be [220, 228-230, 251], even though it is a common, but preventable cause of neonatal brain injury.

Recent studies conducted by the CHYLD group have shown maintaining glucose concentration at or above 2.6 mmol/L may affect neurological function [221], while infants with higher or less stable blood glucose (BG) concentrations in the first 48 hours since birth had higher risk of neurosensory impairment [220]. While research has examined the effect of the level of hypoglycaemia on neurodevelopmental outcomes, it is still unclear what part/s the severity and duration of the dysglycaemia play in these outcomes.

A further study by the CHYLD group sought to investigate factors influencing glycemic stability after neonatal hypoglycaemia and its relationship to neurodevelopmental outcomes [250]. It used many different metrics to quantify glycaemic stability and variability, including time to reach maximum interstitial glucose (IG) concentrations after dextrose boluses, and proportion of BG concentrations outside of a central band in the first 48 hours. The study showed that the risk of neurodevelopmental impairment was increased with both shorter and longer time to reach maximum IG after dextrose boluses, where the middle time to reach maximum IG was between 2.3-4.2 hours. Thus, slow or rapid recovery from hypoglycaemia appears to be associated with neurodevelopmental impairment. However, as discussed in Chapter 6, the metrics used to measure these outcomes can be lacking when trying to accurately and repeatedly quantify glycaemic variability.

As Chapter 6 developed the State Characterisation method, after showing the need for such a metric for variability that takes full advantage of the time series of CGM sensors, the method needs to be validated. This validation may be possible through the association with neurodevelopmental outcomes previously used in the CHYLD study. If quantifying States can predict the outcomes with a high enough sensitivity and specificity, then the algorithm could be seen to accurately capture the patient dynamics and behaviour that is wanted while rejecting the rest of the noise that may confound the links between States and outcomes.

This chapter utilises the State Change algorithm developed in Chapter 6 and [209], to more precisely quantify the glycaemic variability present in this infant cohort. It applies the algorithm with varying parameters to characterise different clinically important glycaemic States, and examines the relationship between these State Changes and the infant cohorts' neurodevelopmental outcomes. The algorithm is generalisable, so it can be adjusted to identify different clinically important behaviours in other cohorts.

7.2 Methods

The State Analysis method presented in Chapter 6 is used again to determine State Changes in an infant cohort. The infant cohort comes from the CHYLD Study in which the association of neonatal glycemia with neurodevelopmental outcomes at 4.5 years [221] and 2 years [220] were assessed. The CHYLD Study recruited 614 infants born from 32 weeks gestation with one or more risk factors for neonatal hypoglycaemia, including the following: diabetic mother; preterm (<37 weeks); small (<10th centile or <2500 g); large (>90th centile or >4500 g); or acute illness [220]. The aim was to examine the relationship between the incidence and severity of neonatal hypoglycaemia in at-risk infants and neurodevelopmental outcome in childhood [220, 229].

A total of 481 of these infants had an interstitial CGMS System Gold CGM sensor (Medtronic Inc., Northridge, CA) inserted soon after birth in the lateral thigh, as previously described [229, 241]. Of these infants, 366 had more than 24 hours of CGM sensor data in the first 48 hours after birth, leading to 12356 total hours of CGM data (median [IQR]: 35.7 [30.5 38.4] hours/patient). Twelve further infants had more than 24 hours of CGM sensor data in the first 48 hours after birth, but had gaps in the data of more than 5 minutes (mean [IQR] = 8.1 [3.0 8.9] hours) and were thus excluded from the analysis. The CGM sensor recorded a measurement every 5 minutes, but results were masked and did not influence clinical care. CGM data were downloaded and recalibrated to all BG concentrations, measured on a blood gas analyser [242].

This cohort is the same one used in the State Change analysis in Chapter 6. However, for the follow up neurodevelopmental assessments, a further 47 infants did not have data for either their 2 year or 4.5 year assessments and so were excluded from further analysis. The resulting total number of infants in this analysis is thus 319.

The State Change algorithm was run a total of 9 times, each time varying the clinical parameters of the algorithm to change the State Changes experienced by each infant. The two parameters varied were:

- Hours allowed between State Changes, where 4 hours, 5 hours and 6 hours allowed were tested. This change tests State length.
- The minimum difference between consecutive State Glycemic Averages, where minimum differences of 0.3, 0.4 and 0.5 mmol/L were tested. This change tests State magnitude in a clinically reasonable range.

These test cases are summarised in Table 7.1. The range of 4-6 hours is based on clinical observations by the CHYLD study principle investigators. The State Change magnitude of 0.3-0.5 mmol/L is also based on their results.

Table 7.1: State Change algorithm parameter testing.

Case Number	Hours Allowed Between State Changes	Minimum difference in State Glycaemic Averages (mmol/L)
1	4	0.3
2	4	0.4
3	4	0.5
4	5	0.3
5	5	0.4
6	5	0.5
7	6	0.3
8	6	0.4
9	6	0.5

After running the algorithm and determining the number of State Changes per day for each patient, this statistic was evaluated against the combined 2 year and 4.5 year neurodevelopmental outcome score. This score provides a simple yes if the infant had any neurodevelopmental problems at either 2 years old or 4.5 years old, or no if there were no problems at either stage. It is based on the combination of a range of clinically accepted scores, as used in [220, 221].

Receiver operator characteristic (ROC) curves where constructed to determine how well the metric of State Changes per Day was able to predict combined neurodevelopmental scores. State Changes per Day was calculated for each patient for standardisation as the hours of CGM recorded for each patient was different.

To construct the ROC curve, the maximum State Changes per Day for a given case was divided by 10 and rounded up to give the maximum number of steps in the ROC curve. The maximum number of State Changes per Day was then set as the condition for neurodevelopmental impairment, and the Sensitivity and Specificity values calculated accordingly. The State Changes per Day threshold for neurodevelopmental impairment was then incremented down in steps of 0.1, recalculating the Sensitivity and Specificity values at each stage, until the threshold was 0.0 State Changes per Day. The

ROC curve was then constructed from the Sensitivity and Specificity values, where each case had a separate ROC curve.

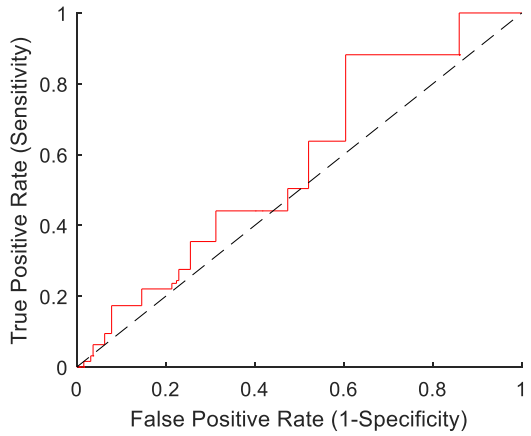
7.3 Results

Table 7.2 presents the results of the State Change algorithm running for the 9 test cases. The number and distribution of State Changes is shown alongside the cohort calculated State Changes per Day median and IQR. In general, results in Table 7.2 show as the hours allowed between State Changes is increased, less States are characterised, as expected. This outcome is also true as the minimum change in State Averages also increases. Increasing both parameters would cause less States to be characterised, so these outcomes make sense logically.

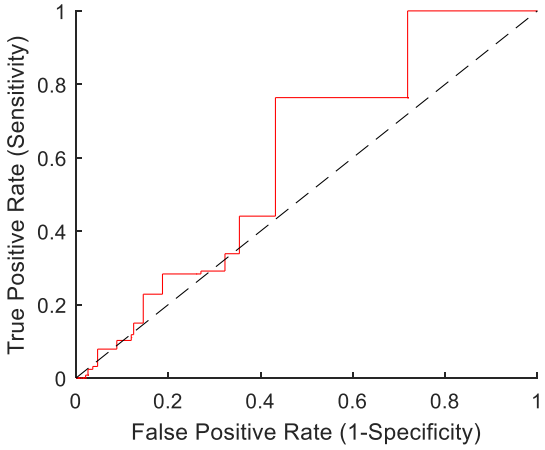
Table 7.2: State Change algorithm parameter testing results.

Case	No State Changes	1 State Change	2 State Changes	3 State Changes	4 State Changes	5 State Changes	State Changes/day (Cohort median [IQR])
1	42	144	76	50	6	1	0.73 [0.55 1.24]
2	84	149	54	29	2	1	0.57 [0.50 1.06]
3	120	134	49	15	1	0	0.54 [0.00 0.80]
4	47	156	70	42	4	0	0.65 [0.54 1.16]
5	88	154	50	27	0	0	0.57 [0.00 1.03]
6	123	137	45	14	0	0	0.54 [0.00 0.71]
7	58	162	70	27	2	0	0.61 [0.53 1.08]
8	98	159	46	16	0	0	0.56 [0.00 0.77]
9	137	138	36	8	0	0	0.53 [0.00 0.63]

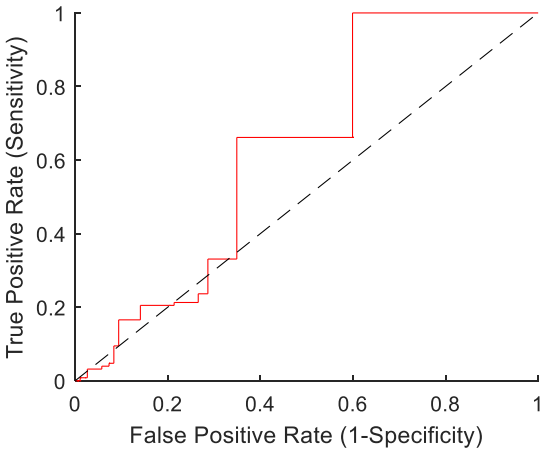
Figure 7.1(a)-(i) show the ROC curves for Cases 1 to 9. The predictive power of State Changes per Day does not appear to be a strong indicator for neurodevelopmental impairment, as seen by the closeness to the diagonal line for all the ROC curves.



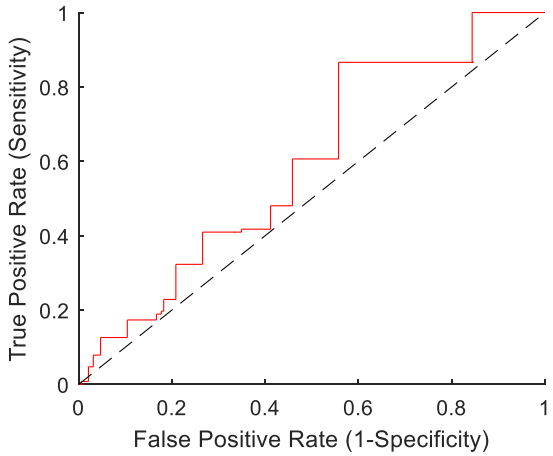
(a) Case 1 - ROC curve for the State Change algorithm run with 4 hours allowed between States, and a minimum 0.3 mmol/L difference in State Averages.



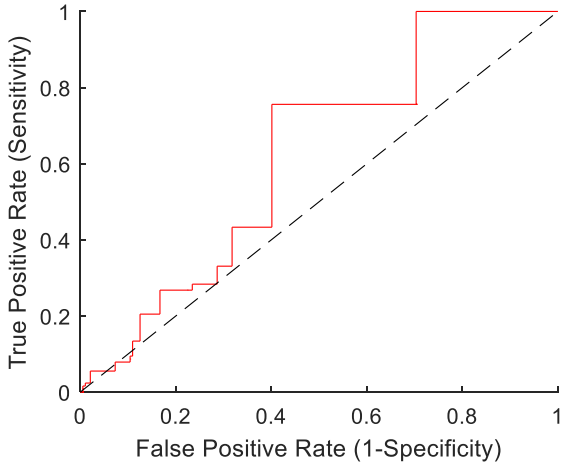
(b) Case 2 - ROC curve for the State Change algorithm run with 4 hours allowed between States, and a minimum 0.4 mmol/L difference in State Averages.



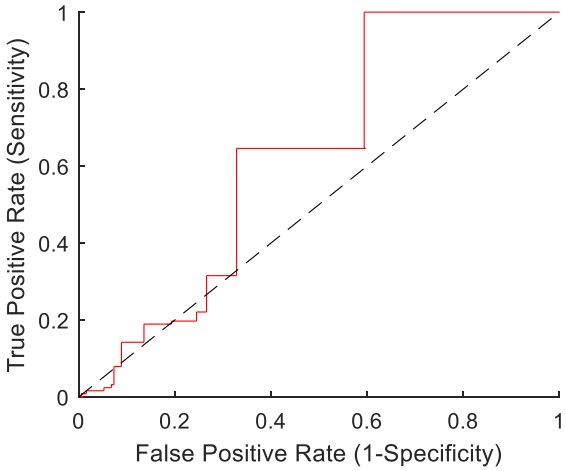
(c) Case 3 - ROC curve for the State Change algorithm run with 4 hours allowed between States, and a minimum 0.5 mmol/L difference in State Averages.



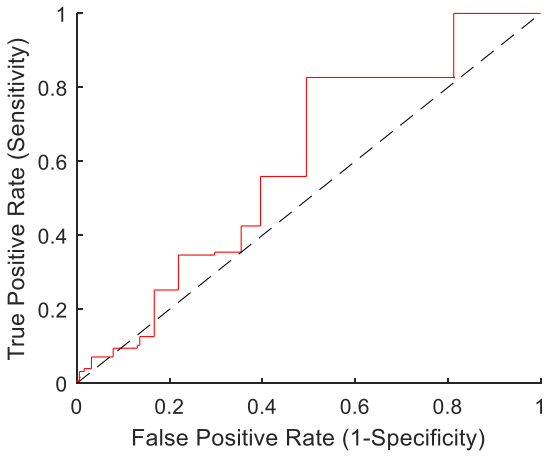
(d) Case 4 - ROC curve for the State Change algorithm run with 5 hours allowed between States, and a minimum 0.3 mmol/L difference in State Averages.



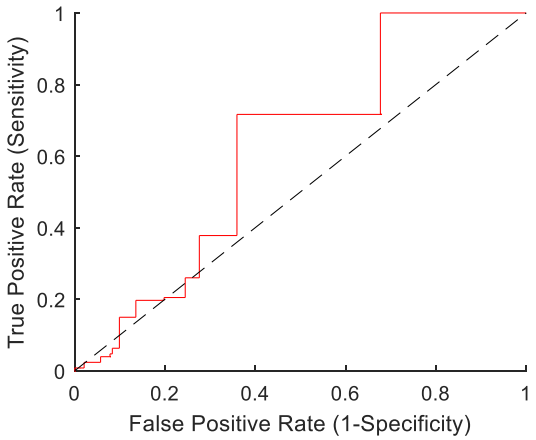
(e) Case 5 - ROC curve for the State Change algorithm run with 5 hours allowed between States, and a minimum 0.4 mmol/L difference in State Averages.



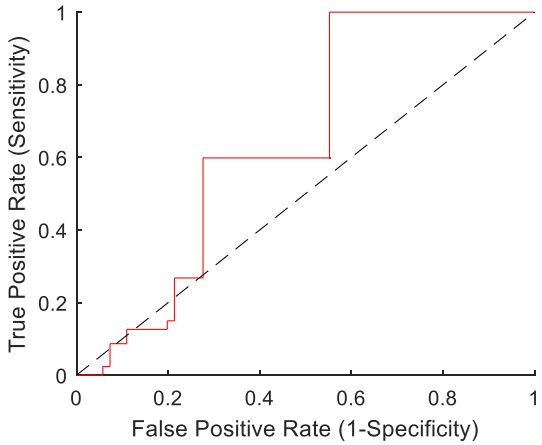
(f) Case 6 - ROC curve for the State Change algorithm run with 5 hours allowed between States, and a minimum 0.5 mmol/L difference in State Averages.



(g) Case 7 - ROC curve for the State Change algorithm run with 6 hours allowed between States, and a minimum 0.3 mmol/L difference in State Averages.



(h) Case 8 - ROC curve for the State Change algorithm run with 6 hours allowed between States, and a minimum 0.4 mmol/L difference in State Averages.



(i) Case 9 - ROC curve for the State Change algorithm run with 6 hours allowed between States, and a minimum 0.5 mmol/L difference in State Averages.

Figure 7.1: ROC curves generated for all State Change algorithm cases run with varying clinical parameters.

7.4 Discussion

Although analysing State Change per Day, and thus glycaemic variability, against neurodevelopmental impairment did not seem to give any conclusive results, it is clear from Burakevych et al. [250] there is some relationship between the glycaemic response to hypoglycaemia, and impairment. The CHYLD Study found slow or rapid recovery of hypoglycaemia was associated with neurosensory impairment. From the State Change analysis, the rapid recovery from hypoglycaemia would appear as a quick State Change, after which could be multiple State Changes if the infant is unable to control their BG levels, or few if they were then stable. In relation to the ROC curves, this outcome would cover the maximum State Changes per Day condition used to predict neurodevelopmental impairment. However, rate of change was not considered here, but could be captured by a shorter average time.

For the slow recovery of hypoglycaemia, the other situational response, this outcome would appear as a long, continuous State. Being glycaemically stable at normal BG levels would usually be considered positive, but to slowly recover and increase BG levels to normal from hypoglycaemia, even after a dextrose bolus, would imply the infant is also struggling to control their BG level, which would be considered detrimental. However, in the case of the State Change analysis and the ROC curves, this outcome would appear as the opposite condition to predict neurodevelopmental impairment, thus giving false negatives and lowering the prediction power, as seen in the plots. There is thus a need to potentially consider State Changes for the duration of BG or similar to capture these effects, which is left for future work.

7.5 Summary

A glycaemic State Change algorithm was applied to an infant cohort to assess glycaemic variability in the first 48 hours since birth. The algorithm was run with differing parameters to capture different levels of clinically important behaviours, and to test for what values were clinically important in this cohort.

ROC curves were formed from each case of parameter testing, using State Changes per Day as the main measurement of glycaemic variability. The max amount of State Changes per Day in any given case was used as the condition for neurodevelopmental impairment in this cohort. The ROC curve analysis was inconclusive, possibly due to the mixed behaviour of recovery from hypoglycaemia that was found to be associated with neurodevelopmental impairment, where a slow or rapid recovery was more detrimental. This slow or rapid recovery from hypoglycaemia would appear as both the minimum and maximum State Changes per Day respectively, and thus multivariate ROC curve analysis may need to be undertaken to find a better association.

Chapter 8 Conclusions

BG dysregulation in a critically ill cohort is associated with increased complications leading to increased morbidity and mortality. Safe and effective external glycaemic control methods have been shown to improve outcomes in this cohort. Such GC protocols include the STAR protocol, currently in use at the Christchurch Hospital ICU in New Zealand. However, achieving safe and effective control consistently, to all patients, has been challenging for most GC methods and protocols. Two major barriers to achieving this consistency are BG measurement frequency and glycaemic variability.

One of the ways BG measurement frequency can be enhanced, without drastically increasing workload, is utilising CGM sensors, which can measure up to 4 times per minute, rather than intermittent measurements, where measurement periods are usually every 1-6 hours, to guide GC protocols. However, using CGM sensors over intermittent measures usually comes with a trade-off of sensor measurement accuracy, which, depending on the severity of the error at a given time, may be detrimental to control. Thus, there is a need to model and characterise these errors, so that CGM sensors can be modelled in-silico to assess the potential trade-offs between safety, performance and workload in GC.

In this work, a CGM sensor was characterised from patient clinical data using an auto-regressive modelling approach. The method presented in this work has the benefits of explicitly accounting for sensor drift and requiring far fewer independently sampled blood glucose measures than other methods. Sensor traces can be simulated for BG taken at a clinically realistic rate to create the model. Sensor simulations showed modelled sensor behaviour was very similar to the original clinical data, with very high similarity in MARD, and equally similar Bland-Altman and Clark Error Grid results further validating the model. The novel use of the Trend Compass to validate the trend accuracy reproduction within simulation further showed that the model method is able to accurately capture both point accuracy and trend accuracy. The overall model method is general to any similar sensor and readily extended to interstitial sensors, with or without including interstitial glucose dynamics. It is easily

simulated on typical clinical data and thus readily able to be incorporated into proven virtual patients to optimise protocol designs to utilise CGMs in the intensive care unit for glycaemic control.

The model of the CGM sensor was then used in place of intermittent BG measurements to guide GC decisions under the STAR protocol in a virtual environment. Virtual trials are used to evaluate the impact on safety, performance, and workload of implementing these or similar CGM sensors for GC using the STAR protocol. Using a range of guard rails and/or rolling windows delineated trade-offs found in using CGM sensors, and decreased the number of required blood draws for BG measures by up to 73%, while also maintaining GC performance and patient safety. Performance and safety were robust to reasonable changes in sensor recalibration period, as well as to changes in rolling window parameters of window length and BG change. Overall, the use of a typical CGM sensor in clinically validated virtual trials shows the potential to reduce clinical workload significantly. The lack of equally significant improvements in performance are likely due to the already very good and clinically proven performance of the STAR GC protocol employed.

Using these characterisation and simulation techniques, this work has shown using CGM sensor technology may improve some aspects of GC, and help the field move towards a future where GC is much more automated. Benefits and areas of improvement can be revealed in a safe in-silico environment, where protocols and treatment parameters can be optimised before transitioning the technology to a real critical care unit.

A further barrier to achieving safe, effective and consistent glycaemic control includes glycaemic variability. Glycaemic variability has been shown to also result in worse outcomes, and so the effective evaluation and control of any variability would help to improve these outcomes. However, current metrics used to identify glycaemic variability are lacking, and do not take advantage of the full time course of data that new CGM sensors can provide.

In this work, the currently used metrics of glycaemic level or State and variability are summarised. Current metrics have significant limitations as most are not clinically defined, and thus poorly represent many aspects needed for a control metric. It thus presents and defines the problem, current state of the art, and presents a vision for the future of what is required, including a first research effort to meet the goals defined. The metric developed uses the full time series data of CGM sensors to characterise different glycaemic states that are clinically important, so that variability can be seen visually.

The overall work includes an ongoing focus to relate the ideas and current state back to traditional control systems methods and engineering approaches. There is still significant room for innovation and new approaches, particularly as more and more data becomes available to engineers, and clinicians become increasingly willing to take on more novel approaches driven by rapid changes in technology and the ability to measure and control patients. These changes will only occur more rapidly, driven in major part by the increasingly poor economics of healthcare mixed with the need to keep care more affordable via increasing automation.

Chapter 9 Future Work

A model for a continuous glucose monitoring sensor was developed to characterise the errors of the Glysure CGM device. The CGM sensor model was implemented in in-silico virtual trials. A new metric to measure glycaemic variability, taking advantage of increased frequency of CGM sensor measurements, was also developed and applied to an infant cohort. Initial results have shown that improvements in workload may be possible, whilst maintaining safety and performance of GC, and that the algorithm can be applied successfully to identify patient glycaemic states. Further avenues of research have been opened up that may build on the work.

9.1 Characterisation of other CGM Sensors for In-silico Testing

So far, this research has only characterised the sensor error of one type of CGM sensor. The characterisation method is generalisable to any CGM sensor, and thus could be used on other data sets to characterise and simulate other sensors in a virtual GC environment, to evaluate the best type of sensor to use for GC.

This is especially important considering the vast number of CGM sensors that are available, many of which utilise different glucose sensing technologies, and also have different insertion sites on the body. Virtual trials could be run, and comparisons made to cheaply and efficiently evaluate the best CGM sensor to use for GC in terms of cost, performance, safety and workload.

9.2 Validation of the State Characterisation Method using Short Term Clinical Outcomes

Before further research is done on relating the State Characterisation Method on long term outcomes such as the neurodevelopmental outcomes in later life in [220, 221], the method needs to be validated more on immediate outcomes in the NICU or ICU. Often, additional data describing patient state is

collected at the time of NICU or ICU admission. These shorter-term diagnoses might be more readily linked to characterised State Changes.

Additionally, States could be characterised for patients undergoing GC. Theoretically, as GC progresses, the patient state should be converging towards a stable state in some specified glycaemic band. This characterisation could potentially be done across a cohort of patients undergoing GC guided by CGM sensor measurements. Collecting additional data such as APACHE scores in the adult ICU may also yield more diagnoses to link State Changes to.

9.3 Further Analysis into the Relationship between Glycaemic States and Neurodevelopmental Outcomes

More research has been carried out analysing the relationship between glycaemic variability during recovery from hypoglycaemia in an infant cohort that has shown that both slow and rapid recovery from hypoglycaemia may be associated with negative outcomes. The State Change algorithm was able to successfully quantify Glycaemic States and variability in this cohort using different clinical parameters for State Changes. However, the work was unable to show any association with neurodevelopmental outcomes using single variable analysis and ROC curves, where the main condition for neurodevelopmental impairment was an excessive amount of State Changes per Day.

Analysis of the State Changes per Day metric using multi-variable analysis, or improving the ROC curve sensitivity and specificity values by evaluating better conditions for neurodevelopmental impairment, would potentially better show how these outcomes are related with glycaemic variability.

9.4 Particle Filters to Improve CGM Sensor Measurement and Accuracy

Utilising CGM sensor measurements to guide GC in an in-silico environment have shown promise in reducing workload, while maintaining safety and performance of GC. However, even with frequent recalibrations, CGM sensors are known to have much higher measurement error than intermittent measures.

To further increase the safety and performance of a GC protocol utilising CGM sensors, a particle filter could be built and tested on in-silico patients. The particle filter would utilise both measurements from the CGM sensor, as well as intermittent measurements from mandatory actions such as recalibrations, to better track blood glucose. The particle filter would be able to combine the accuracy of intermittent measurements with the continuous readings from the CGM sensor to give a more accurate reading of BG than just CGM sensor measurement alone, while also providing the early warning system of any drastic change in BG.

A particle filter would be a better choice of filter over others, such as the Kalman filter, for a few reasons. The particle filter is much more robust to sensor error, can deal with non-Gaussian noise, and can also be used on non-linear models such as the ICING model. Thus, if any filter were to be computationally added to increase the accuracy of the simulated sensor, the recommendation would be to use a particle filter.

References

1. Capes, S.E., et al., *Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview.* Lancet, 2000. **355**(9206): p. 773-8.
2. van den Berghe, G., et al., *Intensive insulin therapy in critically ill patients.* N Engl J Med, 2001. **345**(19): p. 1359-67.
3. Umpierrez, G.E., et al., *Hyperglycemia: an independent marker of in-hospital mortality in patients with undiagnosed diabetes.* J Clin Endocrinol Metab, 2002. **87**(3): p. 978-82.
4. Mizock, B.A., *Alterations in fuel metabolism in critical illness: hyperglycaemia.* Best Pract Res Clin Endocrinol Metab, 2001. **15**(4): p. 533-51.
5. McCowen, K.C., A. Malhotra, and B.R. Bistrian, *Stress-induced hyperglycemia.* Crit Care Clin, 2001. **17**(1): p. 107-24.
6. Krinsley, J.S., *Association between hyperglycemia and increased hospital mortality in a heterogeneous population of critically ill patients.* Mayo Clin Proc, 2003. **78**(12): p. 1471-8.
7. Finney, S.J., et al., *Glucose control and mortality in critically ill patients.* JAMA, 2003. **290**(15): p. 2041-7.
8. Jaattela, A., et al., *Plasma catecholamines in severely injured patients: a prospective study on 45 patients with multiple injuries.* Br J Surg, 1975. **62**(3): p. 177-81.
9. Frayn, K.N., *Hormonal control of metabolism in trauma and sepsis.* Clin Endocrinol (Oxf), 1986. **24**(5): p. 577-99.
10. Chernow, B., et al., *Hormonal responses to graded surgical stress.* Arch Intern Med, 1987. **147**(7): p. 1273-8.
11. Weissman, C., *The metabolic response to stress: an overview and update.* Anesthesiology, 1990. **73**(2): p. 308-27.
12. Christiansen, C., et al., *Hyperglycaemia and mortality in critically ill patients - A prospective study.* Intensive Care Medicine, 2004. **30**(8): p. 1685-1688.
13. Bistrian, B.R., *Hyperglycemia and infection: Which is the chicken and which is the egg?* Journal of Parenteral and Enteral Nutrition, 2001. **25**(4): p. 180-181.
14. Ali, N.A., et al., *Glucose variability and mortality in patients with sepsis.* Crit Care Med, 2008. **36**(8): p. 2316-21.
15. Brunkhorst, F.M., et al., *Intensive insulin therapy and pentastarch resuscitation in severe sepsis.* N Engl J Med, 2008. **358**(2): p. 125-39.
16. Donati, A., et al., *Glycaemic variability, infections and mortality in a medical-surgical intensive care unit.* Crit Care Resusc, 2014. **16**(1): p. 13-23.
17. Grey, N.J. and G.A. Perdrizet, *Reduction of nosocomial infections in the surgical intensive-care unit by strict glycemic control.* Endocr Pract, 2004. **10 Suppl 2**: p. 46-52.
18. Ellahham, S., *Molecular mechanisms of hyperglycemia and cardiovascular-related events in critically ill patients: rationale for the clinical benefits of insulin therapy.* Clin Epidemiol, 2010. **2**: p. 281-8.
19. Aird, W.C., *The role of the endothelium in severe sepsis and multiple organ dysfunction syndrome.* Blood, 2003. **101**(10): p. 3765-3777.
20. Motoyama, T., et al., *Possible role of increased oxidant stress in multiple organ failure after systemic inflammatory response syndrome.* Critical care medicine, 2003. **31**(4): p. 1048-1052.
21. Krinsley, J.S., *Glycemic variability: a strong independent predictor of mortality in critically ill patients.* Crit Care Med, 2008. **36**(11): p. 3008-13.
22. Egi, M., et al., *Variability of blood glucose concentration and short-term mortality in critically ill patients.* Anesthesiology, 2006. **105**(2): p. 244-52.
23. Bagshaw, S.M., et al., *The impact of early hypoglycemia and blood glucose variability on outcome in critical illness.* Crit Care, 2009. **13**(3): p. R91.

24. Pretty, C.G., et al., *Variability of insulin sensitivity during the first 4 days of critical illness: implications for tight glycemic control.* Ann Intensive Care, 2012. **2**(1): p. 17.
25. Signal, M., et al., *Glycemic levels in critically ill patients: are normoglycemia and low variability associated with improved outcomes?* J Diabetes Sci Technol, 2012. **6**(5): p. 1030-7.
26. Penning, S., et al., *Glucose control positively influences patient outcome: A retrospective study.* J Crit Care, 2015. **30**(3): p. 455-9.
27. Kalfon, P., et al., *Severe and multiple hypoglycemic episodes are associated with increased risk of death in ICU patients.* Crit Care, 2015. **19**: p. 153.
28. Hermanides, J., et al., *Hypoglycemia is associated with intensive care unit mortality.* Crit Care Med, 2010. **38**(6): p. 1430-4.
29. Egi, M., et al., *Hypoglycemia and outcome in critically ill patients.* Mayo Clin Proc, 2010. **85**(3): p. 217-24.
30. Chase, J.G., et al., *Improving glycemic control in critically ill patients: personalized care to mimic the endocrine pancreas.* Crit Care, 2018. **22**(1): p. 182.
31. Uyttendaele, V., et al., *Virtual Trials of the NICE-SUGAR Protocol: The Impact on Performance of Protocol and Protocol Compliance.* IFAC-PapersOnLine, 2017. **50**(1): p. 6672-6677.
32. Uyttendaele, V., et al., *Untangling glycaemia and mortality in critical care.* Crit Care, 2017. **21**(1): p. 152.
33. Wintergerst, K.A., et al., *Association of hypoglycemia, hyperglycemia, and glucose variability with morbidity and death in the pediatric intensive care unit.* Pediatrics, 2006. **118**(1): p. 173-9.
34. Stewart, K.W., et al., *Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis.* Ann Intensive Care, 2016. **6**(1): p. 24.
35. Krinsley, J.S., *Effect of an intensive glucose management protocol on the mortality of critically ill adult patients.* Mayo Clin Proc, 2004. **79**(8): p. 992-1000.
36. Chase, J.G., et al., *Implementation and evaluation of the SPRINT protocol for tight glycaemic control in critically ill patients: a clinical practice change.* Crit Care, 2008. **12**(2): p. R49.
37. Preiser, J.C., et al., *A prospective randomised multi-centre controlled trial on tight glucose control by intensive insulin therapy in adult intensive care units: the Glucontrol study.* Intensive Care Med, 2009. **35**(10): p. 1738-48.
38. Kalfon, P., et al., *Tight computerized versus conventional glucose control in the ICU: a randomized controlled trial.* Intensive Care Med, 2014. **40**(2): p. 171-81.
39. Finfer, S., et al., *Intensive versus conventional glucose control in critically ill patients.* N Engl J Med, 2009. **360**(13): p. 1283-97.
40. Hersh, A.M., et al., *Lower Glucose Target Is Associated With Improved 30-Day Mortality in Cardiac and Cardiothoracic Patients.* Chest, 2018.
41. Dubois, J., et al., *Software-guided versus nurse-directed blood glucose control in critically ill patients: the LOGIC-2 multicenter randomized controlled clinical trial.* Critical Care, 2017. **21**(1): p. 212.
42. Wiener, R.S., D.C. Wiener, and R.J. Larson, *Benefits and risks of tight glucose control in critically ill adults: a meta-analysis.* JAMA, 2008. **300**(8): p. 933-44.
43. Kavanagh, B.P. and K.C. McCowen, *Clinical practice. Glycemic control in the ICU.* N Engl J Med, 2010. **363**(26): p. 2540-6.
44. Chase, J.G. and J.L. Dickson, *Traversing the valley of glycemic control despair.* Crit Care, 2017. **21**(1): p. 237.
45. Vanhorebeek, I. and L. Langouche, *Molecular mechanisms behind clinical benefits of intensive insulin therapy during critical illness: glucose versus insulin.* Best Pract Res Clin Anaesthesiol, 2009. **23**(4): p. 449-59.
46. Wissing, K.M., et al., *Prospective randomized study of conversion from tacrolimus to cyclosporine A to improve glucose metabolism in patients with posttransplant diabetes mellitus after renal transplantation.* Am J Transplant, 2018. **18**(7): p. 1726-1734.

47. Van den Berghe, G., *How does blood glucose control with insulin save lives in intensive care?* Journal of Clinical Investigation, 2004. **114**(9): p. 1187-1195.
48. Chase, J.G., et al., *Organ failure and tight glycemic control in the SPRINT study.* Crit Care, 2010. **14**(4): p. R154.
49. Chase, J.G., et al., *Tight glycemic control in critical care--the leading role of insulin sensitivity and patient variability: a review and model-based analysis.* Comput Methods Programs Biomed, 2011. **102**(2): p. 156-71.
50. Chase, J.G., B. Benyo, and T. Desaive, *Glycemic control in the intensive care unit: A control systems perspective.* Annual Reviews in Control, 2019.
51. Uyttendaele, V., et al., *3D Stochastic Modelling of Insulin Sensitivity in STAR: Virtual trials analysis.* IFAC-PapersOnLine, 2018. **51**(27): p. 128-133.
52. Lin, J., et al., *Stochastic modelling of insulin sensitivity variability in critical care.* Biomedical Signal Processing and Control, 2006. **1**(2): p. 229-242.
53. Lin, J., et al., *Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care.* Comput Methods Programs Biomed, 2008. **89**(2): p. 141-52.
54. Evans, A., et al., *Stochastic targeted (STAR) glycemic control: design, safety, and performance.* J Diabetes Sci Technol, 2012. **6**(1): p. 102-15.
55. Hovorka, R., et al., *Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes.* Physiol Meas, 2004. **25**(4): p. 905-20.
56. Plank, J., et al., *Multicentric, randomized, controlled trial to evaluate blood glucose control by the model predictive control algorithm versus routine glucose management protocols in intensive care unit patients.* Diabetes Care, 2006. **29**(2): p. 271-6.
57. Pretty, C.G., et al., *Impact of sensor and measurement timing errors on model-based insulin sensitivity.* Comput Methods Programs Biomed, 2014. **114**(3): p. e79-86.
58. Thomas, F., et al., *Reducing the impact of insulin sensitivity variability on glycaemic outcomes using separate stochastic models within the STAR glycaemic protocol.* Biomed Eng Online, 2014. **13**: p. 43.
59. Pretty, C., et al., *Impact of glucocorticoids on insulin resistance in the critically ill.* Comput Methods Programs Biomed, 2011. **102**(2): p. 172-80.
60. Lonergan, T., et al., *A pilot study of the SPRINT protocol for tight glycemic control in critically Ill patients.* Diabetes Technol Ther, 2006. **8**(4): p. 449-62.
61. Evans, A., et al., *Pilot proof of concept clinical trials of Stochastic Targeted (STAR) glycemic control.* Ann Intensive Care, 2011. **1**: p. 38.
62. Fisk, L.M., et al., *STAR development and protocol comparison.* IEEE Trans Biomed Eng, 2012. **59**(12): p. 3357-64.
63. Stewart, K.W., et al., *Stochastic Model Predictive (STOMP) glycaemic control for the intensive care unit: Development and virtual trial validation.* Biomedical Signal Processing and Control, 2015. **16**: p. 61-67.
64. Penning, S., et al., *Does the achievement of an intermediate glycemic target reduce organ failure and mortality? A post hoc analysis of the Glucontrol trial.* Journal of critical care, 2014. **29**(3): p. 374-379.
65. Krinsley, J.S. and J.C. Preiser, *Time in blood glucose range 70 to 140 mg/dl >80% is strongly associated with increased survival in non-diabetic critically ill adults.* Critical Care, 2015. **19**.
66. Krinsley, J.S., et al., *Diabetic status and the relation of the three domains of glycemic control tomortality in critically ill patients: an international multicenter cohort study.* Critical care, 2013. **17**(2): p. R37.
67. Eslami, S., et al., *Glucose variability measures and their effect on mortality: a systematic review.* Intensive Care Medicine, 2011. **37**(4): p. 583-593.
68. Bland, D.K., et al., *Intensive versus modified conventional control of blood glucose level in medical intensive care patients: a pilot study.* Am J Crit Care, 2005. **14**(5): p. 370-6.

69. Mackenzie, I., et al., *Tight glycaemic control: a survey of intensive care practice in large English hospitals.* Intensive Care Med, 2005. **31**(8): p. 1136.
70. Gartemann, J., et al., *Nurse workload in implementing a tight glycaemic control protocol in a UK hospital: a pilot time-in-motion study.* Nurs Crit Care, 2012. **17**(6): p. 279-84.
71. Aragon, D., *Evaluation of nursing work effort and perceptions about blood glucose testing in tight glycemic control.* Am J Crit Care, 2006. **15**(4): p. 370-7.
72. DiNardo, M.M., M.T. Korytkowski, and L.S. Siminerio, *The importance of normoglycemia in critically ill patients.* Crit Care Nurs Q, 2004. **27**(2): p. 126-34.
73. Schultz, M.J., P.E. Spronk, and H.S. Moeniralam, *Tight glycaemic control: a survey of intensive care practice in the Netherlands.* Intensive Care Med, 2006. **32**(4): p. 618-9; author reply 620-1.
74. Waeschle, R., et al., *Intensive Insulin Therapy on ICU: Comparison of two algorithms to control the blood glucose level.* Intensive Care Med, 2005. **31**(S1): p. S203.
75. DeVries, J.H., *Glucose Variability: Where It Is Important and How to Measure It.* Diabetes, 2013. **62**(5): p. 1405-1408.
76. Rodbard, D., *Glucose Variability: A Review of Clinical Applications and Research Developments.* Diabetes Technol Ther, 2018. **20**(S2): p. S25-S215.
77. Service, F.J., *Glucose variability.* Diabetes, 2013. **62**(5): p. 1398-404.
78. Suh, S. and J.H. Kim, *Glycemic Variability: How Do We Measure It and Why Is It Important?* Diabetes & Metabolism Journal, 2015. **39**(4): p. 273-282.
79. Brunner, R., et al., *Glycemic variability and glucose complexity in critically ill patients: a retrospective analysis of continuous glucose monitoring data.* Crit Care, 2012. **16**(5): p. R175.
80. De Block, C.E., et al., *A comparison of two insulin infusion protocols in the medical intensive care unit by continuous glucose monitoring.* Ann Intensive Care, 2016. **6**(1): p. 115.
81. Hirsch, I.B., *Glycemic variability: it's not just about A1C anymore!* Diabetes technology & therapeutics, 2005. **7**(5): p. 780-783.
82. Monnier, L., et al., *Toward Defining the Threshold Between Low and High Glucose Variability in Diabetes.* Diabetes Care, 2017. **40**(7): p. 832-838.
83. Pickup, J.C., S.C. Freeman, and A.J. Sutton, *Glycaemic control in type 1 diabetes during real time continuous glucose monitoring compared with self monitoring of blood glucose: meta-analysis of randomised controlled trials using individual patient data.* Bmj, 2011. **343**: p. d3805.
84. Salardi, S., et al., *The glucose area under the profiles obtained with continuous glucose monitoring system relationships with HbAlc in pediatric type 1 diabetic patients.* Diabetes Care, 2002. **25**(10): p. 1840-1844.
85. Freckmann, G., et al., *Continuous glucose profiles in healthy subjects under everyday life conditions and after different meals.* Journal of diabetes science and technology, 2007. **1**(5): p. 695-703.
86. McDonnell, C.M., et al., *A novel approach to continuous glucose analysis utilizing glycemic variation.* Diabetes Technol Ther, 2005. **7**(2): p. 253-63.
87. McCall, A.L., et al., *A novel analytical method for assessing glucose variability: Using CGMS in type 1 diabetes mellitus.* Diabetes Technology & Therapeutics, 2006. **8**(6): p. 644-653.
88. Ceriello, A., L. Monnier, and D. Owens, *Glycaemic variability in diabetes: clinical and therapeutic implications.* The Lancet Diabetes & Endocrinology, 2018.
89. Adolfsson, P., et al., *Hypoglycaemia Remains the Key Obstacle to Optimal Glycaemic Control–Continuous Glucose Monitoring is the Solution.* European Endocrinology, 2018. **14**(2): p. 50.
90. Lanspa, M.J., et al., *Percent of time in range 70-139mg/dL is associated with reduced mortality among critically ill patients receiving intravenous insulin infusion.* Chest, 2019.
91. Beck, R., et al., *Validation of Time in Range as an Outcome Measure for Diabetes Clinical Trials.* Diabetes care, 2018.

92. Kovatchev, B.P., et al., *Evaluation of a new measure of blood glucose variability in diabetes.* Diabetes Care, 2006. **29**(11): p. 2433-8.
93. Crane, B.C., et al., *The Development of a Continuous Intravascular Glucose Monitoring Sensor.* J Diabetes Sci Technol, 2015. **9**(4): p. 751-61.
94. Schierenbeck, F., A. Franco-Cereceda, and J. Liska, *Accuracy of 2 Different Continuous Glucose Monitoring Systems in Patients Undergoing Cardiac Surgery.* J Diabetes Sci Technol, 2017. **11**(1): p. 108-116.
95. Wollersheim, T., et al., *Accuracy, reliability, feasibility and nurse acceptance of a subcutaneous continuous glucose management system in critically ill patients: a prospective clinical trial.* Ann Intensive Care, 2016. **6**(1): p. 70.
96. Sechterberger, M.K., et al., *Accuracy of Intra-arterial and Subcutaneous Continuous Glucose Monitoring in Postoperative Cardiac Surgery Patients in the ICU.* J Diabetes Sci Technol, 2015. **9**(3): p. 663-7.
97. Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study, G., et al., *Continuous glucose monitoring and intensive treatment of type 1 diabetes.* N Engl J Med, 2008. **359**(14): p. 1464-76.
98. Kovatchev, B., et al., *Comparison of the numerical and clinical accuracy of four continuous glucose monitors.* Diabetes Care, 2008. **31**(6): p. 1160-4.
99. Akintola, A.A., et al., *Accuracy of Continuous Glucose Monitoring Measurements in Normo-Glycemic Individuals.* PLoS One, 2015. **10**(10): p. e0139973.
100. Beck, R.W., P. Calhoun, and C. Kollman, *Use of continuous glucose monitoring as an outcome measure in clinical trials.* Diabetes Technol Ther, 2012. **14**(10): p. 877-82.
101. Holzinger, U., et al., *ICU-staff education and implementation of an insulin therapy algorithm improve blood glucose control.* Intensive Care Med, 2006. **31**: p. S202.
102. Pretty, C.G., et al., *Hypoglycemia detection in critical care using continuous glucose monitors: an in silico proof of concept analysis.* J Diabetes Sci Technol, 2010. **4**(1): p. 15-24.
103. Chase, J.G., et al., *Impact of human factors on clinical protocol performance: a proposed assessment framework and case examples.* J Diabetes Sci Technol, 2008. **2**(3): p. 409-16.
104. Carayon, P. and A.P. Gurses, *A human factors engineering conceptual framework of nursing workload and patient safety in intensive care units.* Intensive Crit Care Nurs, 2005. **21**(5): p. 284-301.
105. Campion, T.R., et al., *Barriers and facilitators to the use of computer-based intensive insulin therapy.* International Journal of Medical Informatics, 2011. **80**(12): p. 863-871.
106. Wong, X.W., et al., *A novel, model-based insulin and nutrition delivery controller for glycemic regulation in critically ill patients.* Diabetes Technol Ther, 2006. **8**(2): p. 174-90.
107. Zhou, T., J.L. Dickson, and J. Geoffrey Chase, *Autoregressive Modeling of Drift and Random Error to Characterize a Continuous Intravascular Glucose Monitoring Sensor.* J Diabetes Sci Technol, 2018. **12**(1): p. 90-104.
108. Facchinetti, A., et al., *Modeling the glucose sensor error.* IEEE Trans Biomed Eng, 2014. **61**(3): p. 620-9.
109. Facchinetti, A., et al., *Model of glucose sensor error components: identification and assessment for new Dexcom G4 generation devices.* Med Biol Eng Comput, 2015. **53**(12): p. 1259-69.
110. Facchinetti, A., G. Sparacino, and C. Cobelli, *Modeling the error of continuous glucose monitoring sensor data: critical aspects discussed through simulation studies.* J Diabetes Sci Technol, 2010. **4**(1): p. 4-14.
111. Facchinetti, A., et al., *Modeling Transient Disconnections and Compression Artifacts of Continuous Glucose Sensors.* Diabetes Technol Ther, 2016. **18**(4): p. 264-72.
112. Laguna, A.J., et al., *Postprandial performance of Dexcom (R) SEVEN (R) PLUS and Medtronic (R) Paradigm (R) Veo (TM): Modeling and statistical analysis.* Biomedical Signal Processing and Control, 2014. **10**: p. 322-331.

113. Lunn, D.J., C. Wei, and R. Hovorka, *Fitting dynamic models with forcing functions: application to continuous glucose monitoring in insulin therapy.* Stat Med, 2011. **30**(18): p. 2234-50.
114. Rodbard, D., *Characterizing accuracy and precision of glucose sensors and meters.* J Diabetes Sci Technol, 2014. **8**(5): p. 980-5.
115. Bochicchio, G.V., et al., *Multicenter Observational Study of the First-Generation Intravenous Blood Glucose Monitoring System in Hospitalized Patients.* J Diabetes Sci Technol, 2015. **9**(4): p. 739-50.
116. Nohra, E., et al., *Results of a near continuous glucose monitoring technology in surgical intensive care and trauma.* Contemp Clin Trials, 2016. **50**: p. 1-4.
117. Bochicchio, G.V., et al., *Results of a multicenter prospective pivotal trial of the first inline continuous glucose monitor in critically ill patients.* J Trauma Acute Care Surg, 2017. **82**(6): p. 1049-1054.
118. Dungan, K.M., et al., *Determinants of the accuracy of continuous glucose monitoring in non-critically ill patients with heart failure or severe hyperglycemia.* J Diabetes Sci Technol, 2012. **6**(4): p. 884-91.
119. Lee, J.H., et al., *Feasibility of continuous glucose monitoring in critically ill emergency department patients.* J Emerg Med, 2012. **43**(2): p. 251-7.
120. Luijf, Y.M., et al., *Accuracy and reliability of continuous glucose monitoring systems: a head-to-head comparison.* Diabetes Technol Ther, 2013. **15**(8): p. 722-7.
121. Damiano, E.R., et al., *A comparative effectiveness analysis of three continuous glucose monitors: the Navigator, G4 Platinum, and Enlite.* J Diabetes Sci Technol, 2014. **8**(4): p. 699-708.
122. Freckmann, G., et al., *Performance evaluation of three continuous glucose monitoring systems: comparison of six sensors per subject in parallel.* J Diabetes Sci Technol, 2013. **7**(4): p. 842-53.
123. Kosiborod, M., et al., *Performance of the Medtronic Sentrino continuous glucose management (CGM) system in the cardiac intensive care unit.* BMJ Open Diabetes Res Care, 2014. **2**(1): p. e000037.
124. Saur, N.M., et al., *Accuracy of a novel noninvasive transdermal continuous glucose monitor in critically ill patients.* J Diabetes Sci Technol, 2014. **8**(5): p. 945-50.
125. Christiansen, M.P., et al., *A Prospective Multicenter Evaluation of the Accuracy of a Novel Implanted Continuous Glucose Sensor: PRECISE II.* Diabetes Technol Ther, 2018. **20**(3): p. 197-206.
126. Critchell, C.D., et al., *Accuracy of bedside capillary blood glucose measurements in critically ill patients.* Intensive Care Med, 2007. **33**(12): p. 2079-84.
127. Zhou, T., et al., *Continuous Glucose Monitoring Measures Can Be Used for Glycemic Control in the ICU: An In-Silico Study.* J Diabetes Sci Technol, 2017: p. 1932296817738791.
128. Signal, M., et al., *Continuous glucose monitors and the burden of tight glycemic control in critical care: can they cure the time cost?* J Diabetes Sci Technol, 2010. **4**(3): p. 625-35.
129. Holzinger, U., et al., *Real-time continuous glucose monitoring in critically ill patients: a prospective randomized trial.* Diabetes Care, 2010. **33**(3): p. 467-72.
130. Boom, D.T., et al., *Insulin treatment guided by subcutaneous continuous glucose monitoring compared to frequent point-of-care measurement in critically ill patients: a randomized controlled trial.* Crit Care, 2014. **18**(4): p. 453.
131. Goldberg, P.A., et al., *Experience with the continuous glucose monitoring system in a medical intensive care unit.* Diabetes Technol Ther, 2004. **6**(3): p. 339-47.
132. Klonoff, D.C., *Continuous glucose monitoring: roadmap for 21st century diabetes therapy.* Diabetes Care, 2005. **28**(5): p. 1231-9.
133. Deiss, D., et al., *Improved glycemic control in poorly controlled patients with type 1 diabetes using real-time continuous glucose monitoring.* Diabetes Care, 2006. **29**(12): p. 2730-2.
134. Mastrototaro, J.J., *The MiniMed continuous glucose monitoring system.* Diabetes Technol Ther, 2000. **2 Suppl 1**: p. S13-8.

135. Tsalikian, E., et al., *Accuracy of the GlucoWatch G2 Biographer and the continuous glucose monitoring system during hypoglycemia - Experience of the Diabetes Research in Children Network.* Diabetes Care, 2004. **27**(3): p. 722-726.
136. Brunner, R., et al., *Accuracy and reliability of a subcutaneous continuous glucose-monitoring system in critically ill patients.* Crit Care Med, 2011. **39**(4): p. 659-64.
137. Westhoff, D., et al., *Validation and feasibility of two Continuous Glucose Monitoring Systems (CGMS) against point-of-care AccuChek® in critically ill patients; a pilot study.* Netherlands Journal of Critical Care, 2010. **14**: p. 381-7.
138. Rice, M.J. and D.B. Coursin, *Continuous measurement of glucose: facts and challenges.* Anesthesiology, 2012. **116**(1): p. 199-204.
139. Siegelaar, S.E., et al., *Accuracy and reliability of continuous glucose monitoring in the intensive care unit: a head-to-head comparison of two subcutaneous glucose sensors in cardiac surgery patients.* Diabetes Care, 2011. **34**(3): p. e31.
140. Holzinger, U., et al., *Impact of shock requiring norepinephrine on the accuracy and reliability of subcutaneous continuous glucose monitoring.* Intensive Care Med, 2009. **35**(8): p. 1383-9.
141. Boyne, M.S., et al., *Timing of changes in interstitial and venous blood glucose measured with a continuous subcutaneous glucose sensor.* Diabetes, 2003. **52**(11): p. 2790-4.
142. Cengiz, E. and W.V. Tamborlane, *A tale of two compartments: interstitial versus blood glucose monitoring.* Diabetes Technol Ther, 2009. **11 Suppl 1**: p. S11-6.
143. Schierenbeck, F., A. Franco-Cereceda, and J. Liska, *Evaluation of a continuous blood glucose monitoring system using central venous microdialysis.* J Diabetes Sci Technol, 2012. **6**(6): p. 1365-71.
144. Schierenbeck, F., et al., *Evaluation of a continuous blood glucose monitoring system using a central venous catheter with an integrated microdialysis function.* Diabetes Technol Ther, 2013. **15**(1): p. 26-31.
145. Blixt, C., et al., *Continuous on-line glucose measurement by microdialysis in a central vein. A pilot study.* Crit Care, 2013. **17**(3): p. R87.
146. Lin, J., et al., *A physiological Intensive Control Insulin-Nutrition-Glucose (ICING) model validated in critically ill patients.* Comput Methods Programs Biomed, 2011. **102**(2): p. 192-205.
147. Docherty, P.D., et al., *Independent cohort cross-validation of the real-time DISTq estimation of insulin sensitivity.* Comput Methods Programs Biomed, 2011. **102**(2): p. 94-104.
148. Suhaimi, F., et al., *What makes tight glycemic control tight? The impact of variability and nutrition in two clinical studies.* J Diabetes Sci Technol, 2010. **4**(2): p. 284-98.
149. Chase, J.G., et al., *Validation of a model-based virtual trials method for tight glycemic control in intensive care.* Biomed Eng Online, 2010. **9**: p. 84.
150. Dickson, J.L., et al., *Generalisability of a Virtual Trials Method for Glycaemic Control in Intensive Care.* IEEE Trans Biomed Eng, 2017.
151. Chase, J.G., et al., *Next-generation, personalised, model-based critical care medicine: a state-of-the art review of in silico virtual patient models, methods, and cohorts, and how to validation them.* Biomedical engineering online, 2018. **17**(1): p. 24.
152. Chase, J.G., et al., *Model-based insulin and nutrition administration for tight glycaemic control in critical care.* Curr Drug Deliv, 2007. **4**(4): p. 283-96.
153. Breton, M. and B. Kovatchev, *Analysis, modeling, and simulation of the accuracy of continuous glucose sensors.* J Diabetes Sci Technol, 2008. **2**(5): p. 853-62.
154. Chee, F., T. Fernando, and P.V. van Heerden, *Closed-loop glucose control in critically ill patients using continuous glucose monitoring system (CGMS) in real time.* IEEE Trans Inf Technol Biomed, 2003. **7**(1): p. 43-53.
155. Chee, F., et al., *Expert PID control system for blood glucose control in critically ill patients.* IEEE Trans Inf Technol Biomed, 2003. **7**(4): p. 419-25.

156. Le Compte, A.J., et al., *Blood glucose prediction using stochastic modeling in neonatal intensive care.* IEEE Transactions on Biomedical Engineering, 2009. **57**(3): p. 509-518.
157. Dickson, J.L., et al., *External validation and sub-cohort analysis of stochastic forecasting models in NICU cohorts.* Biomedical Signal Processing and Control, 2013. **8**(4): p. 409-419.
158. Dickson, J.L., et al., *Development and optimisation of stochastic targeted (STAR) glycaemic control for pre-term infants in neonatal intensive care.* Biomedical Signal Processing and Control, 2013. **8**(2): p. 215-221.
159. Kovatchev, B., et al., *Feasibility of Long-Term Closed-Loop Control: A Multicenter 6-Month Trial of 24/7 Automated Insulin Delivery.* Diabetes Technol Ther, 2017. **19**(1): p. 18-24.
160. Barnard, K.D., et al., *Closing the Loop in Adults, Children and Adolescents With Suboptimally Controlled Type 1 Diabetes Under Free Living Conditions: A Psychosocial Substudy.* J Diabetes Sci Technol, 2017. **11**(6): p. 1080-1088.
161. Hovorka, R., *Closed-loop insulin delivery: from bench to clinical practice.* Nat Rev Endocrinol, 2011. **7**(7): p. 385-95.
162. Kovatchev, B., et al., *The Artificial Pancreas in 2016: A Digital Treatment Ecosystem for Diabetes.* Diabetes Care, 2016. **39**(7): p. 1123-6.
163. Lewis, D., S. Leibrand, and O. Community, *Real-world use of open source artificial pancreas systems.* Journal of diabetes science and technology, 2016. **10**(6): p. 1411-1411.
164. Reifman, J., et al., *Predictive monitoring for improved management of glucose levels.* J Diabetes Sci Technol, 2007. **1**(4): p. 478-86.
165. Zimmermann, J.B., et al., *Design of a prospective clinical study on the agreement between the Continuous GlucoseMonitor, a novel device for CONTinuous ASSessment of blood GLUcose levels, and the RAPIDLab (R) 1265 blood gas analyser: The CONTASSGLU study.* Bmc Anesthesiology, 2012. **12**.
166. Kuure-Kinsey, M., C.C. Palerm, and B.W. Bequette, *A dual-rate Kalman filter for continuous glucose monitoring.* Conf Proc IEEE Eng Med Biol Soc, 2006. **1**: p. 63-6.
167. Signal, M., et al., *Continuous glucose monitoring and trend accuracy: news about a trend compass.* J Diabetes Sci Technol, 2014. **8**(5): p. 986-97.
168. Bailey, T., et al., *Accuracy of a first-generation intravenous blood glucose monitoring system in subjects with diabetes mellitus: a multicenter study.* J Diabetes Sci Technol, 2013. **7**(6): p. 1484-91.
169. Preiser, J.C., et al., *Near-Continuous Glucose Monitoring Makes Glycemic Control Safer in ICU Patients.* Crit Care Med, 2018. **46**(8): p. 1224-1229.
170. De Block, C.E.M., et al., *Randomized Evaluation of Glycemic Control in the Medical Intensive Care Unit Using Real-Time Continuous Glucose Monitoring (REGIMEN Trial).* Diabetes Technology & Therapeutics, 2015. **17**(12): p. 889-898.
171. Leelarathna, L., et al., *Feasibility of fully automated closed-loop glucose control using continuous subcutaneous glucose measurements in critical illness: a randomized controlled trial.* Crit Care, 2013. **17**(4): p. R159.
172. Kopecky, P., et al., *The Use of Continuous Glucose Monitoring Combined with Computer-Based eMPC Algorithm for Tight Glucose Control in Cardiosurgical ICU.* Biomed Research International, 2013.
173. Preiser, J.C., et al., *Glucose Control in the ICU: A Continuing Story.* J Diabetes Sci Technol, 2016. **10**(6): p. 1372-1381.
174. Krinsley, J.S., et al., *Continuous glucose monitoring in the ICU: clinical considerations and consensus.* Crit Care, 2017. **21**(1): p. 197.
175. Kovatchev, B.P., et al., *In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes.* J Diabetes Sci Technol, 2009. **3**(1): p. 44-55.
176. Patek, S.D., et al., *In silico preclinical trials: methodology and engineering guide to closed-loop control in type 1 diabetes mellitus.* J Diabetes Sci Technol, 2009. **3**(2): p. 269-82.

177. Burmeister, J.J. and M.A. Arnold, *Accuracy of the Ysi Stat Plus Analyzer for Glucose and Lactate.* Analytical Letters, 1995. **28**(4): p. 581-592.
178. Erickson, K.A. and P. Wilding, *Evaluation of a novel point-of-care system, the i-STAT portable clinical analyzer.* Clin Chem, 1993. **39**(2): p. 283-7.
179. Hovorka, R., et al., *A simulation model of glucose regulation in the critically ill.* Physiol Meas, 2008. **29**(8): p. 959-78.
180. Chase, J.G., et al., *Physiological modeling, tight glycemic control, and the ICU clinician: what are models and how can they affect practice?* Ann Intensive Care, 2011. **1**(1): p. 11.
181. Chase, J., T. Desaive, and J.-C. Preiser, *Virtual Patients and Virtual Cohorts: A New Way to Think About the Design and Implementation of Personalized ICU Treatments*, in *Annual Update in Intensive Care and Emergency Medicine 2016*. 2016, Springer. p. 435-448.
182. Thomas, F., et al., *A simple method to model a continuous glucose monitoring signal.* IFAC-PapersOnLine, 2017. **50**(1): p. 8775-8780.
183. Facchinetti, A., et al., *Real-time improvement of continuous glucose monitoring accuracy: the smart sensor concept.* Diabetes Care, 2013. **36**(4): p. 793-800.
184. Keenan, D.B., R. Cartaya, and J.J. Mastrototaro, *Accuracy of a new real-time continuous glucose monitoring algorithm.* J Diabetes Sci Technol, 2010. **4**(1): p. 111-8.
185. Geoffrey, M., R. Brazg, and W. Richard, *FreeStyle Navigator Continuous Glucose Monitoring System with TRUstart algorithm, a 1-hour warm-up time.* J Diabetes Sci Technol, 2011. **5**(1): p. 99-106.
186. Stewart, K., et al., *How Should We Interpret Retrospective Blood Glucose Measurements? Sampling and Interpolation*, in *20th World Congress of the International Federation of Automatic Control*. 2017: Toulouse.
187. Stewart, K.W., et al., *Interpretation of Retrospective BG Measurements.* J Diabetes Sci Technol, 2018. **12**(5): p. 967-975.
188. Stewart, K.W., et al., *Creating smooth SI. B-spline basis function representations of insulin sensitivity.* Biomedical Signal Processing and Control, 2018. **44**: p. 270-278.
189. Zisser, H.C., et al., *Accuracy of the SEVEN continuous glucose monitoring system: comparison with frequently sampled venous glucose measurements.* J Diabetes Sci Technol, 2009. **3**(5): p. 1146-54.
190. Bailey, T., et al., *The Performance and Usability of a Factory-Calibrated Flash Glucose Monitoring System.* Diabetes Technol Ther, 2015. **17**(11): p. 787-94.
191. Gross, T.M., et al., *Performance evaluation of the MiniMed continuous glucose monitoring system during patient home use.* Diabetes Technol Ther, 2000. **2**(1): p. 49-56.
192. Pishko, M.V., *Glucose monitoring by reverse iontophoresis.* Diabetes Technol Ther, 2000. **2**(2): p. 209-10.
193. Furnary, A.P. and Y. Wu, *Clinical effects of hyperglycemia in the cardiac surgery population: the Portland Diabetic Project.* Endocr Pract, 2006. **12 Suppl 3**: p. 22-6.
194. Goldberg, P.A., et al., *Implementation of a safe and effective insulin infusion protocol in a medical intensive care unit.* Diabetes Care, 2004. **27**(2): p. 461-7.
195. van Hooijdonk, R.T., et al., *Point accuracy and reliability of an interstitial continuous glucose-monitoring device in critically ill patients: a prospective study.* Crit Care, 2015. **19**: p. 34.
196. Leopold, J.H., et al., *Point and trend accuracy of a continuous intravenous microdialysis-based glucose-monitoring device in critically ill patients: a prospective study.* Ann Intensive Care, 2016. **6**(1): p. 68.
197. Chase, J.G., et al., *Model-based glycaemic control in critical care-A review of the state of the possible.* Biomedical Signal Processing and Control, 2006. **1**(1): p. 3-21.
198. Amrein, K., et al., *Glucose control in intensive care: usability, efficacy and safety of Space GlucoseControl in two medical European intensive care units.* BMC Endocr Disord, 2014. **14**: p. 62.

199. Chase, J.G., et al., *Integral-based filtering of continuous glucose sensor measurements for glycaemic control in critical care.* Comput Methods Programs Biomed, 2006. **82**(3): p. 238-47.
200. Stewart, K.W., et al., *Nutrition delivery of a model-based ICU glycaemic control system.* Ann Intensive Care, 2018. **8**(1): p. 4.
201. Stewart, K.W., et al., *Nutrition delivery, workload and performance in a model-based ICU glycaemic control system.* Comput Methods Programs Biomed, 2018. **166**: p. 9-18.
202. Hann, C.E., et al., *Integral-based parameter identification for long-term dynamic verification of a glucose-insulin system model.* Computer Methods and Programs in Biomedicine, 2005. **77**(3): p. 259-270.
203. Docherty, P.D., et al., *A graphical method for practical and informative identifiability analyses of physiological models: a case study of insulin kinetics and sensitivity.* Biomed Eng Online, 2011. **10**: p. 39.
204. Docherty, P.D., J.G. Chase, and T. David, *Characterisation of the iterative integral parameter identification method.* Med Biol Eng Comput, 2012. **50**(2): p. 127-34.
205. Lonergan, T., et al., *A simple insulin-nutrition protocol for tight glycemic control in critical illness: development and protocol comparison.* Diabetes Technol Ther, 2006. **8**(2): p. 191-206.
206. Schuster, K.M., et al., *Continuous glucose monitoring in the surgical intensive care unit: concordance with capillary glucose.* J Trauma Acute Care Surg, 2014. **76**(3): p. 798-803.
207. Le Compte, A.J., et al., *Impact of variation in patient response on model-based control of glycaemia in critically ill patients.* 2010.
208. Langouche, L., et al., *Effect of intensive insulin therapy on insulin sensitivity in the critically ill.* The Journal of Clinical Endocrinology & Metabolism, 2007. **92**(10): p. 3890-3897.
209. Zhou, T., J.L. Knopp, and J.G. Chase, *The state of variability: A vision for descriptors of glycaemia.* Annual Reviews in Control, 2019.
210. Finfer, S., et al., *Clinical review: Consensus recommendations on measurement of blood glucose and reporting glycemic control in critically ill adults.* Crit Care, 2013. **17**(3): p. 229.
211. Krinsley, J., *Decreased mortality of critically ill patients with the use of an intensive glycemic management protocol.* Critical Care Medicine, 2003. **31**(12): p. A19-A19.
212. Singer, P., et al., *ESPEN Guidelines on Parenteral Nutrition: intensive care.* Clin Nutr, 2009. **28**(4): p. 387-400.
213. Umpierrez, G.E., et al., *Management of hyperglycemia in hospitalized patients in non-critical care setting: an endocrine society clinical practice guideline.* J Clin Endocrinol Metab, 2012. **97**(1): p. 16-38.
214. Yatabe, T., et al., *The optimal target for acute glycemic control in critically ill patients: a network meta-analysis.* Intensive Care Med, 2017. **43**(1): p. 16-28.
215. Cameron, F., et al., *Statistical hypoglycemia prediction.* J Diabetes Sci Technol, 2008. **2**(4): p. 612-21.
216. Finfer, S., et al., *Hypoglycemia and risk of death in critically ill patients.* N Engl J Med, 2012. **367**(12): p. 1108-18.
217. Cryer, P.E., *Symptoms of hypoglycemia, thresholds for their occurrence, and hypoglycemia unawareness.* Endocrinology and Metabolism Clinics of North America, 1999. **28**(3): p. 495-+.
218. Boyle, P.J., et al., *Plasma-Glucose Concentrations at the Onset of Hypoglycemic Symptoms in Patients with Poorly Controlled Diabetes and in Nondiabetics.* New England Journal of Medicine, 1988. **318**(23): p. 1487-1492.
219. Alsweiler, J.M., C.A. Kuschel, and F.H. Bloomfield, *Survey of the management of neonatal hyperglycaemia in Australasia.* J Paediatr Child Health, 2007. **43**(9): p. 632-5.
220. McKinlay, C.J., et al., *Neonatal Glycemia and Neurodevelopmental Outcomes at 2 Years.* N Engl J Med, 2015. **373**(16): p. 1507-18.
221. McKinlay, C.J.D., et al., *Association of Neonatal Glycemia With Neurodevelopmental Outcomes at 4.5 Years.* JAMA Pediatr, 2017.

222. Cowett, R.M., et al., *Glucose disposal of low birth weight infants: steady state hyperglycemia produced by constant intravenous glucose infusion.* Pediatrics, 1979. **63**(3): p. 389-396.
223. Hall, N., et al., *Hyperglycemia is associated with increased morbidity and mortality rates in neonates with necrotizing enterocolitis.* Journal of pediatric surgery, 2004. **39**(6): p. 898-901.
224. Hays, S.P., E.B. Smith, and A.L. Sunehag, *Hyperglycemia is a risk factor for early death and morbidity in extremely low birth-weight infants.* Pediatrics, 2006. **118**(5): p. 1811-1818.
225. Louik, C., et al., *Risk factors for neonatal hyperglycemia associated with 10% dextrose infusion.* American Journal of Diseases of Children, 1985. **139**(8): p. 783-786.
226. Vaucher, Y.E. and P.D. Walson, *Continuous insulin infusion in hyperglycemic, very low birth weight infants.* Journal of pediatric gastroenterology and nutrition, 1982. **1**(2): p. 211-217.
227. Hey, E. *Hyperglycaemia and the very preterm baby*. in *Seminars in Fetal and Neonatal Medicine*. 2005. Elsevier.
228. Rozance, P.J. and W.W. Hay, Jr., *Describing hypoglycemia--definition or operational threshold?* Early Hum Dev, 2010. **86**(5): p. 275-80.
229. Harris, D.L., P.J. Weston, and J.E. Harding, *Incidence of Neonatal Hypoglycemia in Babies Identified as at Risk.* Journal of Pediatrics, 2012. **161**(5): p. 787-791.
230. Harris, D.L., et al., *Dextrose gel for neonatal hypoglycaemia (the Sugar Babies Study): a randomised, double-blind, placebo-controlled trial.* Lancet, 2013. **382**(9910): p. 2077-83.
231. Srinivasan, G., et al., *Plasma glucose values in normal neonates: a new look.* J Pediatr, 1986. **109**(1): p. 114-7.
232. Diwakar, K.K. and M.V. Sasidhar, *Plasma glucose levels in term infants who are appropriate size for gestation and exclusively breast fed.* Arch Dis Child Fetal Neonatal Ed, 2002. **87**(1): p. F46-8.
233. Hoseth, E., et al., *Blood glucose levels in a population of healthy, breast fed, term infants of appropriate size for gestational age.* Arch Dis Child Fetal Neonatal Ed, 2000. **83**(2): p. F117-9.
234. Cornblath, M., et al., *Controversies regarding definition of neonatal hypoglycemia: suggested operational thresholds.* Pediatrics, 2000. **105**(5): p. 1141-5.
235. Lucas, A., R. Morley, and T.J. Cole, *Adverse neurodevelopmental outcome of moderate neonatal hypoglycaemia.* BMJ, 1988. **297**(6659): p. 1304-8.
236. Griffiths, A.D., *Association of hypoglycaemia with symptoms in the newborn.* Arch Dis Child, 1968. **43**(232): p. 688-94.
237. Griffiths, A.D. and G.M. Bryant, *Assessment of effects of neonatal hypoglycaemia. A study of 41 cases with matched controls.* Arch Dis Child, 1971. **46**(250): p. 819-27.
238. Klonoff, D.C., *A review of continuous glucose monitoring technology.* Diabetes Technol Ther, 2005. **7**(5): p. 770-5.
239. Wallia, A., et al., *Round table discussion on inpatient use of continuous glucose monitoring at the International Hospital Diabetes Meeting.* Journal of diabetes science and technology, 2016. **10**(5): p. 1174-1181.
240. McKinlay, C.J.D., et al., *Continuous glucose monitoring in neonates: a review.* Matern Health Neonatol Perinatol, 2017. **3**: p. 18.
241. Harris, D.L., et al., *Continuous Glucose Monitoring in Newborn Babies at Risk of Hypoglycemia.* Journal of Pediatrics, 2010. **157**(2): p. 198-202.
242. Signal, M., et al., *Impact of retrospective calibration algorithms on hypoglycemia detection in newborn infants using continuous glucose monitoring.* Diabetes Technol Ther, 2012. **14**(10): p. 883-90.
243. Signal, M., et al., *Using Stochastic modelling to identify unusual continuous glucose monitor measurements and behaviour, in newborn infants.* Biomedical Engineering Online, 2012. **11**.
244. Dickson, J.L., C.A. Gunn, and J.G. Chase, *Humans are horribly variable.* Int J Clin Med Imaging, 2014. **1**(2): p. 1-1000142.

245. Tunnell, R.D., B.W. Millar, and G.B. Smith, *The effect of lead time bias on severity of illness scoring, mortality prediction and standardised mortality ratio in intensive care--a pilot study.* Anaesthesia, 1998. **53**(11): p. 1045-53.
246. Peres Bota, D., et al., *The Multiple Organ Dysfunction Score (MODS) versus the Sequential Organ Failure Assessment (SOFA) score in outcome prediction.* Intensive Care Med, 2002. **28**(11): p. 1619-24.
247. Del, C.B., et al., *Severity scores in respiratory intensive care: APACHE II predicted mortality better than SAPS II.* Respiratory care, 1995. **40**(10): p. 1042-1047.
248. Udekwu, P., et al., *Glasgow Coma Scale score, mortality, and functional outcome in head-injured patients.* J Trauma, 2004. **56**(5): p. 1084-9.
249. Shaw, G.M. and J.G. Chase, *Does "treatment failure bias" impact comparisons of ICUs?* Intensive Care Medicine, 2012. **38**(8): p. 1412-1412.
250. Burakevych, N., et al., *Factors influencing glycaemic stability after neonatal hypoglycaemia and relationship to neurodevelopmental outcome.* Scientific reports, 2019. **9**(1): p. 8132.
251. Harding, J.E., et al., *An emerging evidence base for the management of neonatal hypoglycaemia.* Early Hum Dev, 2017. **104**: p. 51-56.